DOCUMENT RESUME

ED 224 835 TM 830 029

AUTHOR	Dorr-Bremme, Don; Catterall, James
TITLE	Costs of Testing. Test Use Project.
INSTITUTION	California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	Nov 82
GRANT	NIE-G-80-0112
NOTE	184p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC08 Plus Postage.
DESCRIPTORS	*Achievement Tests; *Basic Skills; *Cost Estimates; Elementary Education; Interviews; *Program Costs; Psychological Characteristics; Public Education; *School District Spending; Suburban Schools; Teacher Attitudes; *Testing Programs; Test Use; Urban Schools
IDENTIFIERS	*Cost Accounting

ABSTRACT

Phase I of the Test Use Project, begun in 1979, was directed at gaining a representative picture of achievement testing in the nation's public schools. The project was designed to examine testing practices, uses, impacts and costs encompassing a wide range of formal and informal assessment measures. Phase II of the Test Use Project explored the direct and indirect monetary costs as well as the opportunity and psychological costs of testing in an inner city and suburban elementary school, and costs of basic skills testing in their districts. A cost accounting model was selected to identify and determine the magnitude of costs in testing. Data were gathered in the districts from relevant documents, discussions with appropriate officials, and interviews with personnel involved in basic skills testing. School data on costs associated with all achievement testing was collected by formal interviews with principals, instructional staff, school specialists, and resource personnel. Findings by the districts and the two schools regarding testing costs are discussed in separate chapters. A chapter on psychological costs examines teachers' attitudes toward tests and the psychological cost study procedures, and summarizes teacher and student commentaries. Appendices include teacher, administrator and student interview documents. (CM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED224835

Deliverable - November 1982

TEST USE PROJECT
COSTS OF TESTING

Don Dorr-Bremme, Project Co-Director
James Catterall, Project Co-Director

Grant Number
NIE-G-80-0112, P2

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

TM 830 029

CENTER FOR THE STUDY OF EVALUATION
Graduate School of Education
University of California, Los Angeles

20024

The project presented or reported herein was performed pursuant to a grant from the National Institute of Education, Department of Education. However, the opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by the National Institute of Education should be inferred.

ACKNOWLEDGEMENTS

The writing of this report, and completion of the research underlying it, were cooperative efforts.

James Catterall conducted the inquiry on district-level costs and wrote the report of his findings that appears as the second chapter.

Donald Dorr-Bremme, James Burry, Beverly Cabello, and Liza Daniels conducted the case studies at Cityside and Hillview Schools. The data they gathered forms the basis of the third chapter, authored by Dorr-Bremme, as well as the section on psychological costs of testing for teachers written by Burry.

Cabello and Daniels developed the student interviews and authored the section on students' attitudes toward testing.

TABLE OF CONTENTS

	Page
INTRODUCTION	1.1
FINDINGS:	
Costs of Basic Skills Testing in Two Districts	2.1
FINDINGS:	
Costs of Testing in Two Schools	3.1
PSYCHOLOGICAL COSTS:	
Teachers' Attitudes Toward Testing	4A.1
Psychological Cost Study	4B.1
Teacher and Student Commentaries: A Summary	4C.1
APPENDICES	
A. Teacher Interview and Recording Sheet	
B. Administrative Interview & Recording Sheet	
C. Student Interview	

INTRODUCTION

The Test Use Project in Overview

In the 1980's a broad range of issues in educational testing confronts policymakers at all organizational levels. Federal, state and local educational agencies - together with professional and advocacy groups representing educational practitioners, parents and students, and test developers - must address themselves to the implications of diverse and proliferating assessment practices and programs. Helping to inform the decisions that persons in these organizations must make is the goal of the Center for the Study of Evaluation's Test Use Project.

To realize that goal, the Test Use Project is gathering basic information, heretofore lacking, on testing practices, testing's uses and impacts, and testing's costs in public schools across the nation.

The project has taken as its research foci:

-- Achievement testing in reading/English language arts and mathematics.

-- Testing of the latter types as it occurs in public schools at the upper elementary and high school levels, i.e., in grades 4-6 and 10-12.

-- Testing practices, test uses and impacts, and testing costs as manifested within schools.

Test Use Project research has followed from broad definitions of tests and testing. Within the boundaries listed above, the project's inquiry has been designed to encompass a wide range of types of formal assessment measures (e.g., commercially produced norm- and criterion-referenced tests and curriculum-embedded achievement measures; tests

of minimum competency or functional literacy; district-, school-, and teacher-developed tests), as well as less formal assessment techniques (e.g., teacher's observations of interactions with students in class).

The Test Use Project has been conducted in two phases. Research during Phase I was directed at gaining a representative picture of achievement testing in our nation's schools. Phase II, the subject of this report, explores in fuller detail the costs of testing. In the pages which follow, we present first an overview of our Phase I research. The range and breadth of testing uncovered in this survey provides relevant context for considering the costs of testing: it provides a broad outline of the time and effort devoted to testing. Phase II fills in the details about the direct and indirect costs associated with that effort. A description of the design for this study follows.

Research in Phase I

The Test Use Project began in December, 1979. Phase I of the research (lasting two years, from the project's start-up to November 30, 1981) was directed to address three central questions:

1. With what frequency and distribution are particular types of tests given in the upper elementary grades and high school?

2. In what ways do particular types of tests and testing impact on schools and those within them,

   (a) through their very presence, as required or recommended,
   (b) through utilization of their results?

3. What factors influence:

   (a) where and how much particular types of testing are done?
   (b) the ways that types of tests, testing, and test score use impact upon schools and those within them?

A year of planning -- including a literature review, exploratory fieldwork in three school districts, and re-analysis of data from an earlier CSE study of testing (Yeh, 1978) -- led to articulation of these three questions and to a design for survey research that would address them.

To obtain the desired nationally representative picture of testing, we drew a probability sample of 114 school districts stratified on the basis of geographical region, locale, SES, school district size, and minimum competency testing policy. We obtained data from 91 of the selected school districts. The teacher respondents consisted of fourth and sixth grade teachers providing information on their testing practices in reading and math, and tenth grade teachers reporting their testing practices in English or math.

On the basis of the fieldwork interviews and the national survey, the following picture of tests and testing (at the sampled grades and content areas) appeared:

The fourth or sixth grade elementary student is likely to spend about 10 hours a year on reading tests and somewhat more than 12 hours a year on math tests. The tenth grade English student appears to spend more than 26 hours a year on English tests and about 24 hours a year on math tests. These figures include only time for administering tests, but not the time spent preparing for the testing event and scoring, recording, etc. after the test is given. The specific kinds of tests used, as a percentage of the total time devoted to testing in language arts/reading and math, appears in Table 1:

Table 1

Types of Test Used,
As a Percentage of the Total Time
Devoted to Testing

TYPE OF TEST	Elementary Teachers		10th Grade English Teachers	10th Grade Math Teachers
	Reading	Math		
Tests which form part of a statewide assessment program	3	3	3	1
Required Minimum Competency Test	1	2	1	1
Tests included with curriculum materials	28	35	8	17
Other commercially published tests	17	18	6	3
Locally developed and district adopted tests	13	8	5	2
School or teacher developed tests	37	35	74	76

Tables 2 and 3 present the elementary and secondary teacher's responses to questions of how they tend to use the various kinds of assessment devices they administer for different decision-making purposes during the course of the school year. They show that for instructional decision making teachers tend to rely heavily on their own and colleagues' judgment, and on commercial and teacher-constructed curricular measures.

Phase II: Overview to The Costs Study

The goal of the costs study was to obtain an estimate of the direct and indirect monetary costs, as well as the opportunity and psychological costs, of testing in schools and districts.

Table 2

Elementary Teacher Use of Assessment Information for Different Decision-making Purposes
(Percentages reporting use of this information as crucial or important for the specified purpose)

Source/Kind of Information	Planning Teaching at Beginning of School Year		Initial Grouping or Placement of Students		Changing a Student from One Group or Curriculum to Another		Deciding on Students' Report Card Grades	
	Reading	Math	Reading	Math	Reading	Math	Reading	Math
Previous teachers' comments, reports, grades	57	52	62	55	x	x	x	x
Students' standardized test scores	57	54	57	51	55	53	17	16
Students' scores on district continuum or minimum competency tests	51	47	50	45	45	39	20	18
My previous teaching experience	94	94	x	x	x	x	x	x
Results of tests included with curriculum being used	x	x	78	67	83	82	75	77
Results of other special placement tests	x	x	61	56	x	x	x	x
Results of special tests developed or chosen by my school	x	x	x	x	56	52	42	42
Results of tests I make up	x	x	80	86	78	85	92	95
My own observations and students' classroom work	x	x	96	97	99	99	98	99

Table 3

High School Teacher Use of Assessment Information for Different Decision-making Purposes
(Percentages reporting use of this information as crucial or important for the specified purpose)

Source/Kind of Information	Planning Teaching at Beginning of School Year		Initial Grouping or Placement of Students		Changing a Student from One Group or Curriculum to Another		Deciding on Students' Report Card Grades	
	Reading	Math	Reading	Math	Reading	Math	Reading	Math
Previous teachers' comments, reports, grades	28	29	34	40	x	x	x	x
Students' standardized test scores	47	29	49	30	62	39	12	8
Students' scores on district continuum or minimum competency tests	48	30	47	36	53	36	9	5
My previous teaching experience	99	97	x	x	x	x	x	x
Results of tests included with curriculum being used	x	x	45	35	58	43	44	31
Results of other special placement tests	x	x	42	26	x	x	x	x
Results of special tests developed or chosen by my school	x	x	x	x	50	31	28	34
Results of tests I make up	x	x	87	77	92	91	99	99
My own observations and students' classroom work	x	x	99	93	99	97	99	95

Everything that we had discovered in the project to this juncture suggested that considerable on-site work would be needed in a study of testing costs in schools and districts. Specifically, it appeared that ongoing observation and interviewing -- conducted proximal to and focusing on particular assessment events -- would be necessary to enable us to locate and estimate important testing costs.

Early in the planning of the costs study we considered possible frameworks for analyzing the costs of testing. Four major research frameworks were considered: (1) cost accounting, which consists of identifying costs and evaluating their magnitude; (2) cost-effectiveness analysis, which requires examination and evaluation of costs, with benefits measured in units (not necessarily monetary) that are appropriate to the specific testing program under consideration; (3) cost-benefit analysis, which identifies each cost and benefit and then assigns (exclusively) dollar values to each; and (4) an economics of information paradigm, which addresses the matter of the proportion of resources that it is justifiable to spend in the acquisition of information.

Our analyses indicated that the more complex models -- cost effectiveness, cost benefit, and the economics of information paradigm -- did not serve our needs and were innapropriate at this early stage in the development of research on the costs of testing.

A cost-effectiveness analysis would have required that we develop both a measure of the effectiveness of a testing program and a total cost figure expressed in some unit appropriate to the program. But the costs and benefits of testing are multiple and not directly

comparable, and until a single total of costs can be associated with the effectiveness of the test or tests under scrutiny, the model is not strictly applicable.

The limitations, for our purposes, of cost-effectiveness analysis are even further aggravated by the demands of cost-benefit analysis. Cost-benefit analysis would require the incorporation of cost and benefits in exclusively dollar terms. This requirement would apply to all costs, some of which have no conceivable dollar equivalents. Because of this, we did not view cost-benefit analysis as a likely means of yielding useful insights.

The economics of information paradigm would have presented even more practical hurdles than those faced in cost-benefit analysis. In place of converting benefits and costs to dollar equivalents, this model would require each of the benefits and costs to be directly associated with its impact on pupil outcomes, including achievement. Relating elements of testing to schooling outcomes would have been problematic because both the costs and benefits of testing are likely to be difficult to define and their links to pupil outcomes may be remote.

Given the foregoing problems, we chose the cost accounting model for our initial research on testing costs. Through use of this model, our intention was (1) to identify the costs associated with testing for selected schools and districts, and (2) to evaluate the magnitude of costs associated with testing for those selected schools and districts. These are important initial steps, prerequisite to more sophisticated analyses using other paradigms.

Summary of Methods

The Phase II cost study was primarily intended to provide illustrative findings: to yield a comprehensive accounting of the costs of selected types of testing in a very small number of typical schools and districts. To achieve this purpose, and given project resource constraints, we decided to examine the testing costs in two elementary schools and the districts in which they were located.

Given that we had previously collected test-use data in both elementary and high school grades, continuity might suggest that we mount the costs study at these same grade levels. Phase II resources, however, were insufficient to fully examine testing costs at both school grade levels, or even at the high school level alone. Previous project work revealed a much greater variation in testing practices among high school teachers than elementary school teachers. This variation takes the form of differential testing requirements, greater teacher test construction, and marked differences in the form, frequency, and duration of testing events in the high school. Conversely, more required testing appears to occur in elementary schools and teachers devote substantial testing time to instruments accompanying basal curricular series. Our decision therefore was to focus our cost study on elementary school practice.

Two elementary schools were selected for study so as to provide a set meeting the following characteristics:

- two districts and schools which, between them, conduct a full range of types of achievement testing

- two districts and schools which have typical organizational structures and assessment and instructional programs and practices
- two unified school districts which thus include both elementary and secondary schools
- two districts and elementary schools within them which provide a contrast on enrollment size and characteristics of their student population

One of the two schools selected for study was an inner city elementary school which is part of a large metropolitan school district. The student population of this school was comprised predominantly of minority students of lower socioeconomic standing. This school participates in a large number of federal, state, and district special, categorically funded programs, many of which require achievement testing. The second elementary school selected, in contrast, was part of a school district in a small suburban town. This school participates in no categorically funded programs and its student population consist largely of Asian and White, middle class students.

At the district level, data on monetary costs of basic skills achievement testing were collected through examination of relevant district documents and discussions with appropriate district officials. To determine opportunity costs at the district level, interviews were conducted with key personnel involved with activities related to basic skills testing and the use of test results.

At the school level, information on the monetary and opportunity costs associated with all achievement testing was collected via formal, comprehensive interviews with the building principals, instructional staff, and school specialists and resource personnel.

These interviews lasted 1 1/2 - 2 1/2 hours. Supplementary observation of testing in classrooms was also conducted. Both procedures --- the comprehensive interviews and observation of testing events -- were also used to identify the psychological costs of testing for the schools' instructional staffs. Formal student interviews, supplemented by the classroom observations, provided the data base for estimating the psychological costs of testing for students in each school.

In the elementary school in the small suburban district, named Hillview in the following chapters, the building principal was interviewed, as well as all eleven teachers, and the single resource specialist who ran and taught in the school's learning laboratory. Testing event observation was conducted in 2 classrooms at grades 2 and 5, and 10 students from grades 4, 5, and 6 were interviewed.

In the elementary school in the large metropolitan district, called Cityside in this report, the building principal was interviewed, as were 16 teachers, 3 other administrators (special program coordinators), and 2 educational specialists. In addition, observation of actual testing events was carried out in several classrooms, and 10 students each from grades 4, 5, and 6 were interviewed.

FINDINGS: COSTS OF BASIC SKILLS TESTING IN TWO DISTRICTS

In this section we describe the basic skills testing practices in the two districts surveyed. Treating each district in turn, we provide both background information on the districts studied and also the results of our data collection from district offices and schools. We provide a profile of each district and an overview of its basic skills testing program. We then discuss the costs related to the testing program according to our field inquiry. To facilitate comparisons and because of various policy issues that might be informed with these data, we discuss testing costs at the central district level and those incurred district-wide separately before attempting to construct overall cost totals. Following discussions of the two districts, a third section is devoted to our observations and comparisons deriving from both sets of data.

Case I: Littleton District

Littleton District is a small, suburban district which operates 4 elementary schools, a junior high school, and a senior high school. District leaders describe the district organization as highly decentralized and our observations support this: the small central office--two certificated officials plus minimal support staff--occupy the central office, and the six Littleton schools autonomously reach many decisions including some regarding their testing programs. Littleton's community has a relatively stable population, by surrounding area standards, and has witnessed both a typical overall

enrollment loss in recent years and a steady growth in Asian student population. A variety of descriptive data for Littleton are presented in Table 4.

Table 4

Littleton District
[Descriptive Data]

Total Enrollment (1982-83 average daily attendance)		3354 pupils
High School (10 - 12)	1060 pupils	
Junior High School (7 - 9)	915 pupils	
4 Elementary Schools (K - 6)	1379 pupils	
Total Budget		$ 5.6 million
Per Pupil Spending		$ 1836
Other Significant funds		
Title I (Chapter I, ECIA)	$ 40,000	
PL 94-142	$ 40,000	
Percent Minority Pupils (Predominantly Asian) (range is 5% to 50% in elementary schools)		18 %
Number of Teachers		130

Littleton District's Testing Program

Littleton District schools administer a typical array of tests which meet both their own demands for information about their pupils

and also various state mandates which require particular tests at various grade levels. Because of the size of the district, there is no full-time testing coordinator in the central office nor anyone assigned at this level with primary responsiblity for testing. Test coordination is a part-time responsibility of a counselor at the high school and at the junior high, and is one of the principal's responsiblities at the elementary schools. Table 5 summarizes the basic skills testing activities in Littleton District, by type of test and grade level.

Table 5

Summary of Littleton District Basic Skills Testing

Level	Test	Basic Purpose
Elementary	Stanford Achievement Test	Cum records
	SRS Assessment Survey	Cum records
	Grade 4 Proficiency	State Required
	State Assessment (Grades 1,3,6)	State Required
	Metropolitan Achievement Test	Title 1 Evaluation
Junior High	SAT	Counseling/Curriculum review
	Gates MacGintie	Placement
	Metropolitan Math	Placement
	L.A. County Proficiency (7,9)	State Mandate
Senior High	Differential Aptitude Tests	Counseling
	Iowa Test of Educational Development	Curriculum Assessment/ Counseling
	Strong Campbell	Interest Inventory
	Survey of Basic Skills	State Mandate
	Basic Skills Inventory	State Mandate (Required for Graduation)

Table 6
Littleton District Testing Costs in Primary Units
(all units in hours unless otherwise specified)

Central Office Costs

Assistant Superintendent
5% FTE

Coordinator
3% FTE

Secretary
8% FTE

Notes:

1 Administered Fall and Spring.

2 Principal delegates testing at Junior High to counselor.

3 Replacement books. All reused.

4 Pretest/posttest distribution.

5 Scoring services.

6 20 hrs = student conferences
5 hrs = parent communications

7 Scoring & Answer sheets

Totals:

Principal Hours	88
Counselor Hours	261
Clerical Hours	539
Purchase $	3240

Central School Level Costs

ELEMENTARY (K-6)

TEST	Average Per School Principal	Clerical	(Total: 4 schools) Principal	Clerical	Purchases
SAT[1] (1-3)	12	96	48	384	$ 0
State Assess.(3)	9	10	36	40	0
Profic/4 (4)	0	2	0	8	0
Profic/6 (6)	1	0	4	0	0
Totals	33	108	88	432	$ 0

JUNIOR HIGH SCHOOL (7-9)

TEST	COUNSELOR[2]	CLERICAL	PURCHASES
SAT	0	0	
GATES	0	14/1[4]	$ 40[3]
Metro Math	90/60		
Profic.	/60	10/48[4]	$ 1800[5]
Totals	90/120 Pretest/Posttest	24/49	$ 1840

HIGH SCHOOL (10-12)

TEST	COUNSELOR	CLERICAL	PURCHASES
Differential Aptitude Test	4/25[6]	2/4	$ 500[5]
Survey of Basic Skills	3/5	3/5	0
Basic Skills Inventory	4/10	10/10	$ 900[7]
Totals	11/40 Pretest/Posttest	15/19	$ 1400

Table 6 (Continued)
Littleton District Testing Costs in Primary Units
(all units in hours unless otherwise specified)

Classroom Level Costs

ELEMENTARY (K-6)

Test	Hours Per Teacher: Admin.	Other	Total		Number of Classes		Total Hours	Pupil Time Per Pupil
SAT[1] (1-3)	18	12	30	x	24	=	720	18.0 hrs
				+	Lab teacher	=	30 = 750 total SAT	
State Assess. (3)	6.5	8	14.5	x	8	=	116	6.5 hrs
Profic/4 (4)	4	2	6	x	8	=	48	4.0 hrs
Profic/6 (6)	2.5	2	4.5	x	8	=	36	2.5 hrs

JUNIOR HIGH SCHOOL (7-9)

Test	Hours Per Teacher		Number of Classes		Total Hours	Pupil Time Per Pupil
SAT (7,8,9)						
Admin	13	x	50	=	650	13
Pretest	1.5	x	50	=	75	0
Posttest	1.5	x	50	=	75	
					800	
Gates/Mac (7,8,9)						
Admin.	7.5	x	29	=	217.5	7.5
Pretest	minimal			=	10 hrs (total)	0
Posttest	4.5	x	29	=	145 hrs (total)	0
					372.5	
Metro Math						
Admin.	2.5	x	7	=	17.5 hrs	2.5
Pretest	0	x	0	=	0 hrs	0
Postest	1	x	7	=	7 hrs	0
					24.5	
Profic.						
Admin.	9	x	8	=	72 hrs	9
Pretest	1.5	x	9	=	13.5 hrs	0
Posttest	1.5	x	5	=	7.5 hrs	0
					93.0	

HIGH SCHOOL (10-12)

Test	Hours Per Teacher	Pupil Time Per Pupil
Differential Aptitude Test (10)	10	1
Survey of Basic Skills	12	1
Basic Skills Inventory	6	1
	28	

Grand Total Teacher Hours: 2268 hours

Table 7
Littleton District Testing Costs in Dollar Approximations
(Note that this table replicates Table 4 but replaces hour estimates with dollar equivalents)

Central Office Costs

Assistant Superintendent[1]	$ 2000
Coordinator[4]	$ 750
Secretary[2]	$ 1600
Total	$ 4350

Central School Level Costs

ELEMENTARY (K-6)

TEST	Principal[3]	Clerical[2]	Totals
SAT (1-3)	$ 692	$ 3694	$ 4386
State Assess.(3)	519	385	904
Profic/4 (4)	0	77	77
Profic/6 (6)	58	0	58
Totals	$ 1269	$ 4156	$ 5425

JUNIOR HIGH SCHOOL (7-9)

TEST	Counselor[4]	Clerical	Purchases	Total
SAT	$ 0	$ 0	$ 0	$ 0
GATES	0	150	40	190
Metro Math	1803	150	0	1853
Profic.	721	557	1800	3078
Totals	$ 2524	$ 707	$ 1840	$ 5071

HIGH SCHOOL (10-12)

TEST	Counselor	Clerical	Purchases	Total
Differential Aptitude Test	$ 349	$ 58	$ 500	$ 907
Survey of Basic Skills	96	77	0	173
Basic Skills Inventory	168	192	900	1260
Totals	$ 613	$ 327	$ 1400	$ 2340

Notes:

[1] Based on $ 40,000 salary and fringes

[2] Based on $ 20,000 salary and fringes

[3] Based on $ 30,000 salary and fringes

[4] Based on $ 25,000 salary and fringes

Totals: Principals $ 1269 Clerical $ 5190
Counselors $ 3137 Purchases $ 3240

Total School Central Level Costs: $ 12,836

Table 7 (Continued)
Littleton District Testing Costs in Dollar Approximations

Classroom Level Costs

ELEMENTARY (K-6)

TEST	Teacher Cost
SAT[1] (1-3)	$ 7776
State Assess. (3)	1253
Profic/4 (4)	518
Profic/6 (6)	389
Total	$ 10,260*

JUNIOR HIGH (7-9)

TEST	Teacher Cost
SAT (7,8,9)	$ 8640
Gates/Mac (7,8,9)	4023
Metro Math	265
Profic.	1004
Total	$ 13,932

SENIOR HIGH (10-12)

TEST	Teacher Cost
Differential Aptitude	$ 108
Survey of Basic Skills	130
Basic Skills Inventory	65
Total	$ 302

* Rounding error not reconciled.

Totals:
Cost of Teacher Time: $ 24,494

The Costs of Testing in Littleton District

We investigated the costs of the various basic skills assessments conducted by Littleton District during the school year 1981-82. The methods of the investigation were outlined in detail in a previous section of this report, but an overview of their important elements may be useful to the reader at this point.

The principal tasks of this phase of our research were to identify the various ingredients of the basic skills testing activities of the district, to attain estimates of the magnitude of each of these costs in their primary units (such as teacher or counselor hours devoted to testing, or direct dollar costs of materials and services purchased), and finally to convert all resource estimates to dollar equivalents. The rationale for this approach flows simply from the potential uses for information revealed in our research about testing costs. From a decision-making standpoint, the overall level of resources committed to basic skills testing has meaning when compared to the total of resources available to the district for all of its operations. And from instructional and service standpoints, the time devoted to testing by pupils, teachers, counselors, administrators, and support staff may be important in the context of the overall allocation of time among tasks for district personnel.

We began by interviewing district personnel at all levels to identify the types of tests administered and the full range of district resources attached to their basic skills testing. We probed the nature of test administration, pre-test and post-test activities of personnel, various analysis and dissemination activities at the

classroom, school, and central office levels, and the types of materials and services purchased from outside vendors. After achieving a satisfactory idea of what seemed to be involved in Littleton's testing, we surveyed district personnel at all levels to generate estimates of dollars expended or time involved in testing activities. Key respondents were one of Littleton's two assistant superintendents, his secretary, the principal of each school, the counselors in charge of testing at the junior high and high schools, and the teachers themselves.

Table 4 presents a summary of the types of costs identified, and the actual estimates for each of these costs in their primary units. These data can inform a host of questions which we will not attempt to catalogue here, but a few examples may help to illustrate the substance and organization of the information.

It is apparent from the central office presentation that basic skills testing is not a major activity at this level in Littleton District, since it occupies between 3 percent and 8 percent of work time for these individuals. Data reflecting this are shown in Table 6 as fractions of time spent on all testing matters by three individuals at the central level--the assistant superintendent, a program coordinator, and a clerical staffer. None of the respondents was able to suggest a finer breakdown of his time than this, such as significant allocations to one particular test or to testing at particular grade levels. We were reminded by these respondents that the administrators of individual schools were chiefly responsible for all testing functions in their domains.

The central school-level costs display in Table 6 refers to those testing costs above the classroom level at the six schools in the district. At the elementary schools, these costs are for the time of principals and clerical staff at each school; at the junior high, test coordination is the responsibility of a counselor who is assisted by clerical staff, and in addition some dollar costs for scoring services and materials were identified for junior high testing; at the high school level, counselor time, clerical staff time, and material and service purchases were identified, and the personnel hours involved are reported accordingly in the table.

The classroom level costs reported in Table 6 include the hours devoted to testing by teachers, and the amount of pupil time spent in testing by each pupil in the district. One apparent fact of Littleton basic skills testing from this display is that time spent in district-mandated, basic-skills testing appears to be rather negligible at the high school levels in comparison to the earlier grades. This is reflected in much lower totals of both teacher hours and pupil hours devoted to testing.

Additional observations drawn regarding the information in Table 6 (and from the dollar estimates contained in Table 7) will be presented below. We will first describe the conversion of our various personnel time estimates into dollar cost estimates as the second step in our analysis of district testing costs.

Table 7 replicates Table 6 with one important difference: where Table 6 showed the number of hours devoted to testing by a variety of district personnel, Table 7 converts each of these estimates to dollar equivalents. This is done by applying estimated annual personnel cost figures for each category of staff involved in testing (teachers,

principals, administrators, counselors, and clerical staff), and then estimating the value of the time devoted to testing by each as an appropriate share of their annual cost to the district. The annual cost estimates for each personnel classification appear as notes in Table 7, and were drawn to include fringe benefits and other direct employee costs beyond typical salaries. Table 7 thus presents dollar estimates for the costs of each test, at each level, and affords some detail in showing just where these costs occur. For instance, the SAT test in the elementary schools commands the personnel resources of principals ($692), clerical staff ($3694), and teachers ($7776). This can be contrasted with the 4th grade proficiency test which engages comparatively few resources in its administration and handling (clerical costs of $77 and teacher costs of $518). Many similar comparisons can be drawn with these data.

Pupil time shown in Table 6 has not been converted to dollar estimates, although there are conceivable purposes for such an activity. The pupils do not engage fractions of the district's budget in the manner of other personnel involved in district activities, and therefore do not represent direct or indirect costs to the district that have a meaningful dollar interpretation. Nevertheless, as we cited in the theoretical development of our testing cost inquiry, the amount of time spent by pupils in various activities can be thought of as having various costs and benefits, particularly those accruing to the effectiveness of the instructional programs of the district. Pupil time estimates from this study may have value in secondary analyses or related research, but are not featured in the present analysis.

We suggested that Tables 6 and 7 lend themselves to a variety of analyses that may be of interest to a cost of testing inquiry. The next displays summarize the cost data of Table 7 in several ways. They attend to broad questions such as comparisons of testing costs to overall spending in the district, the degree to which testing costs are incurred as a result of outside mandates for assessments, and how pupil time is spent in testing at each level.

Table 8
Littleton Testing: Costs Per Pupil, and Cost Summary, by Level

Level	Total Monetary Costs			Total Costs Per Tested Pupil	Costs Per Pupil At Level
	Central	Teacher	Total		
Central Office	$ 4350		$ 4350		$ 1.30 per pupil
Elementary					
SAT*	$ 4386	$ 7776	$ 12162	$ 20.27	
State Assess.*	904	1253	2157	3.60	
Prof 4*	77	518	595	2.98	
Prof 6*	58	389	447	2.24	
All Tests	$ 5425	10260	$ 16132		$ 11.70 per Elementary pupil
Junior High					
SAT	$ 0	$ 8640	$ 8640	$ 28.33	
Gates	190	4023	4213	4.61	
Metro	1803	265	2068	3.39	
Prof.*	3708	1004	4082	6.60	
All Tests	$ 5071	$ 13932	$ 19003		$ 20.77 per Junior high pupil
High School					
DAT	$ 907	$ 108	$ 1015	$ 2.88	
SBS*	173	130	303	0.86	
BSI*	1260	65	1325	1.25	
All Tests	$ 2340	$ 303	$ 2643		$ 2.49 per High school pupil

* State mandates

Table 8 summarizes the dollar cost estimates from Table 7, and shows the magnitude of these costs in per-pupil terms. The costs per pupil tested for each test and at each level are shown immediately to the right of the dollar totals. These costs range from a high of $28.33 for the SAT test at the junior high to a low of $0.86 for the SBS test at the high school. In addition, the total costs of testing per pupil enrolled at each level are shown at the extreme right of Table 8. The central office resources devoted to testing translate to $1.30 per pupil districtwide. The junior high devotes the most resources to testing ($20.77 per pupil), and this amount is just about one percent of the district's average per pupil expenditure ($1836 per pupil). Overall, it appears that Littleton testing costs amount to about one half of one percent of the overall total of district expenditures.

Table 9

Littleton District: Direct vs. Indirect Cost of Basic Skills Testing, by Level

Level	Testing Costs Per Pupil	Direct Share	Indirect Share
Central Office	$ 1.30		100%
Elementary	$ 11.70	negligible	100%
Junior High	$ 20.77	9.7%	90.3%
High School	$ 2.49	53%	47%

Table 9 shows what fraction of the testing costs per pupil at each level in Littleton can be accounted for by direct versus indirect costs. For this purpose, we have included as direct costs those items for which the district incurs an expenditure of funds, such as the

cost of test booklets, answer sheets, and scoring services. The indirect costs represent the share of personnel time (or its dollar equivalent) devoted to testing activities. With the exception of the high school testing, it appears that the vast majority of testing costs are bound up in the time of district personnel who administer the tests and who analyze and disseminate the results. In contrast, the high school testing program experiences realtively high direct costs since the activities occupy comparatively few teachers, who are needed for few hours, and at the same time incur comparatively high costs for scoring services.

Table 10

Littleton District: Mandated vs. District Discretionary Testing Costs, by Level

Level	Overall Basic Skills Testing Costs Per Pupil	Mandate Share	Discretionary Share
Elementary	$ 11.70	24.6%	75.4%
Junior High	$ 20.77	21.5%	78.5%
High School	$ 2.49	61.6%	38.4%

Some tests administered in Littleton result from the district's own decisions about assessment needs, while others must be administered to satisfy state requirements. Table 10 shows the share of testing costs at each of the elementary, junior high, and high school levels resulting from each of these two types of tests. Again, a contrast is apparent between the high school and lower levels. About a fourth of Littleton testing below grade 10 is done in response to outside mandates, while more than half of the costs of testing in the high school are tied directly to such mandates.

Summary Comments: Littleton District Testing Costs

As we stated earlier, the overall cost accounting for test costs in Littleton could inform a variety of questions, many of which are not raised here explicitly. Issues of who is involved in testing (principals versus counselors versus support staff), or issues of which types of tests seem to incur which type of costs are examples of such supplementary inquiries. We highlight here a few overall observations that stand out as we examine this profile of Littleton's testing costs.

First, the central office testing costs are minimal-- equivalent to about a dollar per pupil. As we will see in our discussion of a much larger district subsequently in this report, this has some consistency with what we found to be true when a great number of central resources (multiple staff, scoring, and purchases) are devoted to the testing of a large number of students. Second, the magnitude of testing costs overall is small in comparison to overall resource expenditure in the district, on the order of a half a percent of total district expenditures. And within this small total cost for testing, a generally small fraction is accounted for by direct dollar expenditures for such things as tests, materials, and scoring. As such, from a budgetary standpoint, Littleton's testing occupies a nearly negligible portion of its total resources, and of those costs that are attributable to testing, by far the most important are the costs of teacher and administrator time devoted to the process. This suggests to us that the dollar costs of testing may be less important than other considerations attached to the personnel time that generates most of those costs, such as effective use of teacher or principal time. Overall, it appears that the testing "budget" per se,

"budget" per se, even in the broadest sense of including personnel time allocations, is not a potential gold-mine should Littleton seek resources for other endeavors.

Overview of Metro District

Metro District is a major urban school district with most of the characteristics attendant to that identity. The pupil population is diverse, the district maintains hundreds of schools and employs thousands of teachers, and the district budget is a complex mix of general support and state and federal categorical programs aimed at specific types of pupils. Table 11 highlights some of Metro District's dimensions that are of interest to our study.

Table 11
Metro District: Descriptive Data

Total Enrollment (1981-82)		543,791
High School (10-12)	127, 221 pupils	
Junior High School (7-9)	120,337 pupils	
Elementary School (K-6)	291,632 pupils	
Schools for Handicapped	4,601 pupils	
Total Budget		$ 1.84 billion
Per pupil spending includes:		
Basic State Aid per pupil	$ 1,890	
Local revenues per pupil	409	
Federal Programs per pupil	330	
State Categoricals per pupil	320	
Other Revenues per pupil	351	
Student Racial/Ethnic Composition		
American Indian	0.37%	
Asian/Pacific Islander	7.5 %	
Black	22.2 %	
Hispanic	47.4 %	
White	22.5 %	
Number of Schools		
Elementary	427	
Junior High Schools	75	
High Schools	49	
Magnet Schools/Centers	84	

Number of Classroom Teachers	Total	Average Per Grade Level
Elementary	9721	1389
Junior High	3539	1180
High School	3742	1247

Metro District spends nearly twice as much money annually per pupil on average than Littleton District, but about all of this difference is accounted for by the presence of specially funded programs. The district pupil population is largely non-white, with significant representation from several minority groups.

Metro District's Testing Program

As we found with Littleton, Metro District conducts a variety of basic skills tests for a variety of internal and external purposes. The tests administered, at which grade levels, and for which reasons are outlined in summary form in Table 12. The largest single testing effort is the skills test given to all children in grades 1 through 6, the Continuum-Based Skills Survey (CBSS). This test was developed by the district and its consultants over a several year period and is used primarily so that teachers will have good information about the performance of children in their classes. The test also satisfies state and federal reporting requirements for Chapter I, ECIA (formerly Title I, ESEA) program for grades 3 and 5.

Other tests and their purposes are also listed in Table 12. Beyond the CBSS test, these are dominated by the grade 7 and grade 10 proficiency assessments which are given to students initially at these levels, and repeatedly (if necessary) until they are passed. Three tests--one each for math, writing, and reading-- are administered for these proficiency assessments at each level. The high school assessment is conducted in response to a state mandate which requires districts to establish such testing as a requirement for graduation. The junior high proficiency tests represent a district decision to assure pupil performance prior to high school entry, although pupils

may enter 10th grade without having passed the junior high battery of proficiency tests. Finally, some of Metro District's testing, is done to satisfy reporting requirements for federal and state aid programs. The CTBS is administered to fulfill these requirements at various levels in addition to the administration of the CBSS test in grades 3 and 5 which doubles for district and federal purposes.

Table 12
Metro District: Overview of Basic Skills Testing

	Test	Grades	Type	Rationale
Elementary	CBSS	1-6	Criterion-referenced	Pupil diagnosis, curriculum planning, 3-5: Chapter I reports to State/Fed
	CTBS	3,5 (6 optional)	Norm-referenced	Instructional program assessment.
	CTBS Espanol	1-6	Spanish version	Individual tests for all children receiving Spanish reading instruction.
	CAP	entry,1,3,6		State Assessment
Junior High	ASC	7 plus retakes	Proficiency	Pupil progress, math
	Writing Profic.	7 plus retakes	Proficiency	Pupil progress, language, writing
	PAIR	7 plus retakes	Proficiency	Pupil progress, reading
	CTBS	8	Norm-referenced	Instructional program assessment.
	CTBS (Chapter I schools)	7,8,9	Norm-referenced	State/Federal reports.
Senior High	Math Profic.	10 plus retakes	Proficiency	H.S. graduation requirement-math.
	Writing Profic.	10 plus retakes	Proficiency	H.S. graduation requirement-writing, language
	READ Sr.	10 plus retakes	Proficiency	H.S. graduation requirement-reading
	CTBS (Chapter I schools)	10-12	Norm-referenced	State/Federal reports (10 out of 49 schools)

Metro District Central Office Testing Costs

The size and organization of Metro District dictate a somewhat different approach to the assessment of district testing costs from the one pursued in Littleton and reported above. The guiding questions are the same: what is the full range of elements which constitute the costs of conducting basic skills testing in the district? Which tests are accompanied by which types of costs? What is the magnitude of these costs? And what is the importance of these costs from the standpoint of overal district resource management? But since there are hundreds of schools and thousands of teachers and other individuals involved in the process, our research necessarily could not take as microscopic a look at testing activities as we were able to in the case of a much smaller district.

The first problem we faced in this very large district was the fact that testing responsibilities lay in many offices throughout the district, and that no one person has a complete view of the full array of testing practices and related activities. The second, another problem that we anticipated, was that the various officials charged with administration of testing were not accustomed to thinking about the various costs of what they oversee. The district does not budget for testing in ways that correspond to the types of questions in our interest. We were therefore presented with a substantial and formative schedule of detective work, and the results left us with a great many partial perspectives of the objects of our inquiry. What follows is a report of our attempts to reconcile these views onto an overall ledger.

In contrast to the smaller Littleton, Metro District assigns significant central resources to its basic skills testing programs, both in the form of personnel who administer and coordinate the testing programs, and in direct purchases of processing services and materials. The central office houses five professional and five clerical staff who work exclusively with district tests. One professional oversees the entire testing program, one administers Chapter I (compensatory education) testing programs, and the other three divide up responsibility for the remaining tests administered. The activities of these individuals have largely predictible descriptions--scheduling tests and all related activities, coordinating purchase and delivery of materials, arranging for test scoring, writing reports of test results, and ongoing development of the testing programs.

District testing coordinators also conduct inservice training of field personnel including principals, coordinators of testing at the school level, and area directors of instruction. The inservice training schedule is heaviest in October and January when 2 to 3 day-long sessions per week are customarily scheduled and conducted by one or more of the 5 central office coordinators.

The central office also houses two automated scoring machines which are used whenever machine scorable answer sheets accompany tests. These machines require a total of between 4 and 6 operator handlers when tests are being scored. In addition, the central office requires the services of about two full-time equivalent computer programmer/consultants to assist in its information processing needs for scoring and information handling.

Table 13

Metro District: Central Costs Not Specific to Particular Tests
($ in 1000's)

Job Identification	Number FTE	Annual Cost ($1000)
Basic Skills		
Professional/coordinator	4.1	$ 150
Clerical	4.0	80
Compensatory Education		
Professional/coordinator	1.0	35
Programmers	1.9	65
Clerical	1.0	20
Scanning		
Operator/handlers	5.0	100
Programmer/consultant	.2	7
		$ 457
Office Space		$ 10
Transportation		10
Warehousing		5
		$ 25
Total Central Office		$ 482
Total Cost per pupil		$ 0.89

Table 13 summarizes the costs incurred by Metro District to maintain its central testing related services. These costs are predominantly found in the various personnel allocated to testing in the central office. The total central cost, $ 482,000, represents a cost of just under one dollar per pupil enrolled in Metro District.

In addition to maintaining a central coordination and administration staff for its basic skills testing, Metro District incurs significant central costs for testing through a variety of services and purchases outside of the central office which

nevertheless remain above and beyond any costs incurred in the schools themselves. These costs are summarized in Table 14.

Table 14

Metro District: Summary of Annual Costs
Above School Level, Outside Central Office

Cost	Amount ($1000)
Development of CBSS	$ 120
Area Scoring Centers	$ 400
Supplies	$ 120
Test Processing and Handling	$ 103
Contract Scoring	$ 211
Total	$ 954
Average cost per pupil	$ 1.75

The most significant cost of the testing program outside of the central office costs is the operation and maintenance of the area scoring centers in the district's 10 regional offices. The 1981-82 estimate of these costs was $400 thousand which is allocated primarily to "seasonal" employees who are hired temporarily during peak times of test scoring. (This arrangement is being changed for the coming year to one in which a certificated professional at each site will have full responsibility for area scoring center activities. Overall costs will not be affected by this change.) In addition Metro District contracts with vendors outside of the immediate central district office for test processing and handling. Supply costs for all tests (booklets, answer sheets, pencils) are estimated to total $120 thousand annually. Finally Metro District has entered into a long term contract with an outside laboratory for the development of its elementary skills assessment CBSS test. The cost of this service in

1981-82 was about $120 thousand (it has gone down each year), and the total spent for this contract since its inception since 1976 is about $1 million.

The total cost of these additional services and purchases ($954 thousand) represents about $1.75 per pupil district wide in Metro District. The grand total of testing costs in Metro District which occur above the school level ($1.436 million) represents about $2.64 per pupil enrolled in the district. These estimates are highlighted in Table 15.

Table 15

Total Metro District Testing Costs Above the School Level
(all $ amounts in 1000's)

Central Office Costs	$ 482
Other Central Costs	$ 954
Total	$ 1,436
Average cost per pupil	$ 2.64

The Costs of Specific Testing Conducted in Metro District

Costs incurred by Metro District for each of its basic skills tests are shown in Table 16. These figures represent a mixture of direct budgeted costs revealed to us in internal district documents, the estimated costs of personnel assigned to functions attached to specific tests, and the pro-rating of costs of central testing functions that are not specifically attributable to any one particular test or group of tests. The direct costs for materials and contract scoring are maintained in district accounting records. Estimates of processing and handling costs were obtained from the same records.

The allocation of area scoring center costs was achieved through estimates obtained in interviews of share-of-activity devoted to the various tests. District office personnel costs were assigned on the basis of reported share of personnel time devoted to specific tests. The remaining costs of testing ($307 thousand) were allocated across tests according to the number of pupils actually tested in each assessment during the school year.

Table 16
Metro District: Central Costs by Test

	DIRECT COSTS				DISTRICT OFFICE				
Test	Materials	Contract Scoring	Processing & Handling	Area Scoring Center	Profess.	Clerical	Share of Unallocated Costs[1]	Contract Development	TOTALS
CBSS	$ 5	$ 0	$ 15	$ 200	$ 19	$ 10	$ 98	$ 120	$ 467
CTBS	3	0	0	50	19	10	80	—	162
ASC	20	83	5	25	12.5	7	21	—	173.5
Writing Proficiency (Jr. High)	teacher graded		36	25	12.5	7	21	—	101.5
READ Jr.	60	45	11	25	12.5	7	21	—	181.5
Math Proficiency	20	83	6	25	12.5	7	22	—	175.5
Writing Proficiency (Sr. High)	6	teacher graded	30	25	12.5	7	22	—	102.5
READ Sr.	6	0	0	25	12.5	7	22	—	72.5
	120	$ 211	$ 103	$ 400	$ 113	$ 62	$ 307	$ 120	$ 1436

[1] Based on share of total pupils tested for each test.

The total testing costs for each test are again displayed in Table 17, along with per pupil testing costs for each test.

Table 17

Metro District: Costs of testing Per Pupil Tested, by Test

TEST	TOTAL COSTS	COSTS PER PUPIL TESTED[1]
CBSS	$ 467	$ 1.60
CTBS	162	1.55
ASC	173.5	3.50
Writing Proficiency (Junior High)	101.5	2.03
READ Jr.	181.5	3.63
Math Proficiency	175.5	2.93
Writing Proficiency (Senior High)	102.5	1.71
READ Sr.	72.5	1.21
	$ 1436	$ 2.64

[1] Numbers of pupils tested estimated using enrollments by grade level, plus estimates of test retakes for proficiency tests.

School Level Testing Costs, Metro District

We now turn to the costs of testing in Metro District that lay beyond the district's central office. Recall that we consulted with personnel who coordinate testing at the district's central office and achieved an overall estimate suggesting that Metro District spends about $2.64 per pupil for these activities. Here we investigate testing costs incurred in the schools themselves, including those involving administrators, counselors, coordinators, and secretaries as well as the teachers who administer most tests.

Because of limitations in our investigative resources, we have not generated what can be presented as a representative view of the Metro District's more than 500 regular schools, so what follows is merely a suggestion of what the cost patterns would look like if certain similarities were to obtain between what we observed and the testing practices in the balance of the district's schools. At the elementary level, we conducted an exhaustive study of the testing costs in a "typical" Metro District school (Cityside) which are reported in the next chapter. We extend these findings across all of the district's elementary schools to estimate the total of resources devoted to testing at this level. At the junior high and high school levels, we do not even have limited field work to draw from. (Recall that project resources precluded fieldwork at the secondary level.) For projected total costs at the secondary level, we examine what we learned abut testing costs in our other study district (Littleton), and calculate what must be considered to be, at best, illustrative figures for the much larger Metro District. At both the elementary and secondary levels, we use information derived in our national survey of test use to suggest what types of tests may account for the costs we do identify.

Elementary Testing Costs

Our extensive case study of the Cityside Elementary school in Metro District afforded us a rich view of its various costs related to testing of all types conducted during the 1981-82 school year. These were reported in Table 30 in this volume, and this distribution is incorporated into Table 18 below which projects these cost findings across the remainder of the district's elementary schools.

Table 18 shows our case study findings regarding the central office costs as well as the direct and indirect costs to schools of conducting all testing over the 1981-82 school year. These tests include basic skills tests (of the sort we investigated in-depth for the Littleton District), and also include the various tests that teachers use solely for curricular or pupil progress assessments. Column (A) presents the costs for all contributing personnel, services, and materials in per-pupil terms. The cost per pupil at Cityside school for all testing activities is estimated to total $130, or less than 7 percent of the district's total general expenditures per pupil.

Estimates of the total cost of testing across the district's 427 elementary schools, which are displayed in column (C) of Table 18, were calculated by means of a linear extrapolation from what we observed in the case study. The projected grand total of testing costs for Metro District elementary school is about $38 million, which

TABLE 18

Estimates of Total METRO DISTRICT Elementary Level Testing Costs
Per Cityside School Case Study

TYPE OF COSTS			(A) Total at Cityside [Enrollment = 830]	(B) Per Pupil Cost	(C) Estimated Total Costs All Elementary Schools [Enrollment = 291,000]
District-Office Costs[1]:					
$2.64 per pupil x 830 pupils			$ 2191	$ 2.64	$ 768,000
Direct Costs to School:					
Purchase of Metropolitan Achievement Test			1200		
Purchase of Curricular Reading Tests			6000		
Purchase of Scantron Scoring Machine Forms			200		
			$ 6400	$ 7.71	$ 2,244,000
Indirect Costs for School (Personnel Time):	Hours/Year(% Work Time)[2]		Dollar Equivalents[3]		
Administrators/Coordinators -					
Reading Resource Teacher	328.5 (19.3%)		$ 5790		
Title I Program Coordinator	11.5 (0.7%)		210		
Teacher Testing Coordinator	35.0 (2.1%)		472		
	375.0		$ 6472	$ 7.80	$ 2,270,000
Clerical/Secretarial	10.3 (0.5%)		$ 95	$ 0.11	$ 32,000
Classroom Teachers -					
Average Time Per Teacher	199.2 (12.2%)		$ 2745		
Number of Teachers	x 30		x30		
	5975.32		$ 82,350	$99.22	$28,934,000
Instructional Specialists[4] -					
Bilingual Coordinator	156.25 (9.2%)		$ 2760		
Bilingual Teacher (assists with testing)	8.08 (0.5%)		112		
	164.33		2872	$ 3.46	$ 1,007,000
Instructional Aides (Paraprofessionals) -					
Aide to Reading Resource Teacher (n = 1)	109.45(20.6%)		$ 657		
Aide to Instructional Specialist (n = 1)	4.58 (0.9%)		$ 27		
Classroom Aides (per classroom)	39.48 (7.8%)	$ 237			
Number of Classrooms	x 30	x 30			
	1184.50		$ 7110		
	298.5				
Classroom Volunteers	92.2(??)		7794	$ 9.39	$ 2,732,000
Student Time[5] -					
Average Time Per Pupil	76.1 (8.6%)		----		
TOTAL COSTS FOR SCHOOL (1981-82 School Year)			$ 108,174		$37,987,000 (or about $130 per pupil)
AVERAGE COSTS PER CLASSROOM (n = 30; avg 27.67 pupils/class)			$ 3606		
COSTS PER PUPIL			$ 130.33	$ 130.33	
PROPORTION OF DISTRICT ANNUAL EXPENDITURE PER CHILD (= $1890)			6.9%		

represents about 6.9 percent of the district's total per pupil expenditures. However, we would expect actual total per pupil expenditures in a unified school district to be less at the elementary level than at the secondary level. (The more elaborate nature of school programs at upper levels makes them more costly.) Therefore, the actual share of costs at the elementary level attributed to testing is probably higher than this 6.9 percentage estimate.

Table 19

Distribution of Total Costs for Testing Per Pupil
in Metro District: Elementary Grades by Type of Test
[Per Cityside Case & Per National Survey Estimates of Distribution]

Type of Test	Distribution Per Case[1]		Distribution Per National Survey[2]	
	%	$	%	$
State Assessment	7.0%	$ 9.09	3.0%	$ 3.91
MCT's			1.5%	1.95
Curriculum Materials Tests	38.1%	49.66	31.5%	41.06
Other, Commercially Published	8.3%	10.82	17.5%	22.81
Locally Developed	3.3%	4.30	10.5%	13.68
School or Teacher Developed	43.3%	56.46	36.0%	46.92
	100.0%	$ 130.33	100.0%	$ 130.33

1 Dorr-Bremme, Table E, Table C

2 Choppin, B. "How Schools Make Use of Test Results" Center for the Study of Evaluation. Revised April 1982. Table 4.

Both our Cityside School case study and the national survey of testing practices in the schools allows us to estimate what types of tests account for the more than $130 worth of resources per pupil estimated to be devoted to testing in Metro District's elementary schools. According to our respondents at Cityside School, the vast majority of these resources are devoted to tests imbedded in curriculum materials or to tests developed by teachers or the schools themselves. Table 19 shows that more than 80 percent of testing resources are directed toward these tests (commercial curricular plus teacher developed tests). The data further show that only about 7 percent of testing resources are expended to satisfy state requirements for pupil assessment and demonstration of competencies. Table 19 also shows that the reported distribution of testing resources at Cityside School does not depart radically from national patterns of test use at the elementary level.

Junior High and High School Testing Costs

Our reports of total Metro District costs for testing at the secondary level do not benefit from an empirical excursion into these schools (we could not conduct one). It is, rather, a sketch of what cost patterns might look like if what we found in our analysis of Littleton District applied in the much larger Metro District. We present these calculations as being simply illustrative, and without further analysis of the 100+ secondary schools in Metro District, we have no basis for claiming that the dollar figures reported truly reflect resources expended for testing at this level. This portrayal

of school level costs at the secondary level in the Metro District is further hampered by the fact that our Littleton District analysis surveyed only basic skills testing and not testing done to satisfy curriculum requirements. So the analysis which follows is restricted to basic skills testing at the secondary level, which typically accounts for considerably less than half of all testing activity.

The analytical reasoning we employ below is straight forward. If per pupil costs for basic skills testing at the Metro District junior high and high schools are equivalent to what we observed in Littleton, the total basic skills testing costs in the much larger Metro District may be obtained by simple multiplication of the per pupil cost estimates by actual enrollments. Furthermore, if these costs are incurred in similar patterns in both districts across the different types of resources used in testing (chiefly the costs of various personnel and materials), we can base the estimated distribution of Metro District costs on the pattern observed in Littleton. And, in addition, our national survey of testing practices at the secondary level allows us to suggest just which types of tests these resources might be devoted to. We now proceed with these constructions, despite their limited foundations.

Table 20.1

Projected Basic Skills Testing Costs in Metro District:
Junior High School

[Based on Littleton District Estimates of School Level
Costs & Metro District Central Cost Analysis]

	Cost By Category	Total Metro District Costs [120,000 Enrollment]
Central Cost*	$ 2.64	$ 316,800
Administrators/ Counselors	2.67**	320,400
Clerical	0.75**	90,000
Teachers	14.71**	1,765,200
	$ 20.77 per pupil**	$2,492,400
	(< 1% of district jr. high budget per pupil)	(< 1% of district jr. high budget)

* Estimated in Metro District Central Office Analysis. Includes Purchases of Materials/Services.

** Derived from Tables 5 and 6.

As shown in Table 20.1, if the $20.77 overall per pupil cost for basic skills testing in Littleton were to characterize Metro District costs for the same activities, the district would spend a total of about $2.4 million on these tests in its junior high schools. This represents a little less than 1 percent of the average per pupil general expenditure districtwide. If the distribution of these costs is also similar to that observed in the smaller district, where the costs of teacher time account for about three-fourths of the basic skills testing resources, this $2.5 million would be distributed as shown in the right-hand column of Table 20.2.

Table 20.2

Distribution of Metro District Junior High Basic Skills Testing Costs:
[Per Total Cost Estimates (Table 19) and
National Survey of Test Use Distributions.]

Type of Basic Skills Test	% of All Basic Skills Test Time Reported[1]	Per-Pupil Cost Distribution
State Assessment	29%	$ 6.02
MCT	6%	1.25
Local or District Developed	29%	6.02
Other, Commercially Developed	36%	7.48
		$ 20.77 per pupil

[1] Choppin, op. cit; based, on 10th grade observations.

Our national survey of testing practices suggests that different types of basic skills tests might occupy differing amounts of time at the junior high school level.* Table 20.2 incorporates the distribution of basic skills type tests observed nationally, and displays the application of this distribution to the $20.77 in per pupil resources we have identified as suggestive of Metro District junior high test costs. As we have previously pointed out, about a third of all basic skills testing at this level is done to satisfy state mandates, and the balance is intended to satisfy local demand for basic skills development information.

* Our 10th grade estimates from the survey are used for these projections. No junior high grades were surveyed.

Table 21

Projected Basic Skills Testing Costs in Metro District: High Schools
[Based on Littleton District Estimates of School Level Costs]

	Cost By Category	Total Metro District Costs [127,000 Enrollment]
Central Cost*	$ 2.64	$ 335,300
Administrators/ Counselors	0.59**	74,900
Clerical	0.31**	39,400
Teachers	0.28**	35,600
	$ 3.82 per pupil**	$ 485,200
	(< 1% of district budget per pupil)	(< 1% of district budget)

* Estimated in Metro District Central Office Analysis. Includes Purchases of Materials/Services.

** Derived from Tables 5 and 6.

Table 21 and Table 22 present treatments analogous to those presented for junior high school estimates in order to derive estimates for Metro District high school level basic skills testing costs. Littleton District reported "spending" only $3.82 per pupil for basic skills testing efforts in their junior high schools. A similar level of costs in the Metro District would imply a total of about half a million dollars would be devoted to basic skills testing for the 127,000 pupils in its high schools (Table 21). The pattern of costs among resources (shown in the same table) is weighted comparatively toward administrators and counselors at the high school

level. Littleton reported a predominance of centrally administered basic skills tests, and the distribution shown here reflects their comparative underuse of teachers for test administration. The total cost of basic skills testing in the Metro District high schools suggested this presentation would amount to a small fraction of one percent of the district's budget.

Table 22 shows how this small level of testing costs at Metro District high schools would be allocated across different types of basic skills tests, if the patterns were similar to those found in our national survey of schools. In comparison to the junior highs, these costs are somewhat more tied to state assessments and competency testing, but are still dominated by local demands for basic skills testing.

Table 22

Distribution of Metro District High Schools Basic Skills Testing Costs

[Per Total Cost Estimates (Table 18) and National Survey of Test Use Distributions.]

Type of Basic Skills Test	% of All Basic Skills Test Time Reported[1]	Cost Distribution
State Assessment	14%	$ 0.53
MCT	14%	0.53
Local or District Developed	29%	1.11
Other, Commercially Developed	43%	1.65
		$ 3.82 per pupil

[1] Choppin op. cit.; based on 10th grade observations.

As we stated at the outset of this discussion of testing costs within Metro District's schools, our limited efforts to gain a representative view of the more than 500 elementary and secondary schools in the district severely restrict our ability to provide concrete estimates of what is actually spent on testing by Metro District beyond the central office level. In Littleton District, we were able with simple surveys and interviews to capture a relatively complete portrait of district testing practice. The sheer size of the Metro District, with its great diversity of schools and pupils, demands a research budget beyond the one at our disposal if achieving reliable total cost estimates is the target. So what we have presented in this section, and specifically the information contained in Tables 18 through 22, is a characterization of school level testing costs which is based on a very partial view of actual practice in the district, on inferences drawn from our in-depth study of a smaller district, and on our national survey of testing practices.

FINDINGS: THE COSTS OF TESTING IN TWO SCHOOLS

The preceding section has provided an accounting of basic-skills testing costs in the Littleton and Metro School Districts. Now, focus shifts to the costs of testing in one elementary school in each of these districts. The following pages provide a detailed look at the costs of all achievement testing in these schools in the basic skills but also in other subject areas.

As noted in the introduction, information for these cost accountings was gathered in extended interviews with the school's administrators, classroom teachers, and instructional specialists. They were asked to describe the time and other resources that they and their students expended on achievement testing of all types in all school subjects through the 1981-82 academic year. The interviews were conducted in May and June of that year, with some follow-up during September to clarify details and confirm data. (Refer to the introduction of the research methodology.)

Testing Costs in Littleton District's Hillview School

Hillview is the smallest of Littleton's four elementary schools. Its eleven classrooms and learning laboratory serve 191 students: 50% of Asian background, about 45% from White Anglo families, the remaining 5% Hispanic or Black. Specific socioeconomic indices were unavailable, but the neighborhood from which Hillview children come is considered one of the higher-income areas in generally well-to-do Littleton. Homes within the school's attendance boundaries are valued in the $250,000 - $400,000 range, substantially above the $120,000 average for the county. Students' parents work largely in professional, executive, and scientific-research positions.

Hillview participates in no special, educational programs sponsored by the state or federal government. Its program is supported exclusively by Littleton District funds.

The school has a reputation for excellence in the Littleton District, and its students are considered "very high achievers" by the teaching staff. As the principal noted, "A so-called "average" kid (in terms of national norms) is not average here. He's below average."

Hillview educators are experienced, and most have been at the school for some time. The principal has served at Hillview for fifteen of his twenty-six years as a head administrator. The teachers' length of service at Hillview is, on the average, nine years. Most taught elsewhere before joining the Hillview faculty.

To present a comprehensive summary of Hillview's testing program is difficult; there is considerable variation from classroom to classroom. Table 23, however, presents an overview of those measures that are widely and/or consistently administered. In addition to those shown are various tests and quizzes developed or selected by individual faculty members. (A fuller picture of the scope of Hillview's achievement assessment will emerge during the following discussion.)

The foregoing has been a brief introduction to Hillview Elementary School and its testing program. An accounting of testing costs at Hillview follows.

Hillview Testing Costs in Overview

Table 24 itemizes the total costs for all achievement testing reported for Hillview during the 1981-82 school year. Most entries in

TABLE 23

Hillview Elementary School Testing Program

Test	Grade(s)	Required by:	Administrations Per Year
Multi-Subject			
Stanford Achievement Test	K - 6	District	2
Otis-Lennon Intelligence Test	K - 6	District	2
State Assessment Program	1,3,6	State	1
Reading			
Ginn 720 Placement Test	1*	District	1
Ginn 720 Criterion (Unit) Test	1 - 6	District	9-20†
Ginn 720 Mastery Test	1 - 6	District	1-2
Ginn 720 Booster Test	1 - 6	District	As needed
Math			
Scott-Foresman Unit Pre-Test	2 - 6	District	5-12†
Scott-Foresman Unit Post-Test	1 - 6	District	5-12
District-Developed MATH Operations Test	1 - 5	District	Weekly-monthly
Math Proficiency Test	4	District	1
Junior High School Math Placement	6	District	1
Spelling			
Teacher-Developed or Commercial-Curriculum Spelling Test	1 - 6		Bi-weekly or weekly
Physical Education			
Physical Performance Test	5	State	1

* The instructional specialist in the Hillview learning laboratory also routinely administers the Ginn placement test to all students new to the District except those not proficient in English.

† Variations noted in the frequency of curricular testing were reported from classroom to classroom. In some instances, variations ocured within classrooms where individualization of instruction permitted learners to progress through the curriculum at different rates.

this table are self-explanatory, especially in light of the accounting procedures employed and explained in the previous chapter. Derivations of the "present work time" and the dollar equivalents for staff time are clarified in footnotes to the table.

The first item, district-office costs, is incurred in the time personnel in Littleton District's Central Office devote to testing. (See Tables 7 and 8 in the foregoing chapter.) Here, the $1.30 per pupil cost is applied to Hillview's 191 students.

As is the case with other Littleton elementary schools, Hillview makes no direct purchases in conjunction with testing. The district and state supply various mandated tests. Consumable test booklets that accompany commercial curriculum materials in reading and math are bought by the district. (In the district budget, these costs are included under general outlays for instructional materials. They could not be differentiated and pro-rated for Hillview. A rough estimate, however, suggests that the cost of these curriculum-embedded testing materials would be under $1,000 for Hillview's 191 students.)

Of course, teachers consume paper, duplicating fluid, ditto masters, and even chalk in the process of producing their own tests. But no one at Hillview would venture to estimate what proportion of these and similar supplies went for testing. In any case, the cost of routine stationery supplies for testing is almost certainly minimal.

Table 24 makes apparent, then, that virtually all of Hillview Elementary School's economic testing costs are indirect: i.e., they are the dollar values of the staff time devoted to testing. As indirect dollar costs they are borne by the district, which pays staff salaries. But the staff time invested in testing can also be

TABLE 24

Total Costs for All Achievement Testing in
HILLVIEW SCHOOL - LITTLETON DISTRICT
(Enrollment = 191)

	Hours/Year(% Work Time)[2]	Dollar Equivalents[3]
District Office Costs[1]:		
$1.30 per pupil x 191 pupils		$ 248
Direct Costs to School:		None reported
Indirect Costs for School (Personnel Time):		
Administrators/Coordinators -		
Principal	63.75(3.75%)	$ 1125
Teacher Testing Coordinator	36.00(2.12%)	477
	99.75	$ 1602
Clerical/Secretarial		None reported
Classroom Teachers -		
Average Time Per Teacher Number of Teachers	252.96(15.5%) x 11	$ 3875 x 11
	2782.50	$ 42,625
Instructional Specialists[4] -		
Learning Laboratory/English as a Second Language	197.63 (11.6%)	$ 2510
Instructional Aides (Paraprofessionals) -		None employed
Classroom Volunteers	77.66 (??)	——
Student Time[5] -		
Average Time Per Pupil	88.04 (9.9%)	——
TOTAL COSTS FOR SCHOOL (1981-82 School Year)		$ 47,085
AVERAGE COSTS PER CLASSROOM (n = 11; avg 17.36 pupils/class)		$ 4280.45
COSTS PER PUPIL		$ 246.52
PROPORTION OF DISTRICT ANNUAL EXPENDITURE PER CHILD (= $1836)		13.4%

[1] Calculations of District Office Costs are Shown in Chapter Two

[2] The "% Work Time" figures are based on respondents' report of hours worked per week before, during, and after school hours. These reported hour per week were averaged by role category across the two schools studied (Cityside and Hillview). Reported hours were within similar ranges at both schools. Work times used are as follows:

(a) For administrators, coordinators, and instructional specialists: 46 hours per week x 37 weeks per year.

(b) For classroom teachers: 44 hours per week x 37 weeks per year = 1628 hours per year.

(c) No total hours per unit or person could be ascertained for volunteers.

[3] Dollar equivalents are based upon the proportion of work time expended at the following salary estimates:

(a) For administrators and coordinators - $ 30,000 salary and fringe benefits

(b) For classroom teachers and the instructional specialist - $ 22,500 salary and fringe benefits.

These salary estimates are equivalent to those used in the analysis of District costs, but are 20% - 25% lower than those actually in effect in this school.

[4] Instructional specialist time reported is devoted to assessing the language competence of incoming students, other placement testing of new students, and recurrent assessment of students enrolled in an English as a Second Language (ESL) course.

[5] Student time shown equals the time spent by the typical student in each classroom averaged across the school's regular classrooms. The percentage shown is based on 5 class hours per day (not counting the hour for lunch and recess) for 177 school days per year, which equals 885 classroom hours per school.

construed as an opportunity cost -- that is, as the allocation of a resource to one activity (testing) instead of another (for example, explicit instruction). Seen from this perspective, the cost of testing in staff time is borne by multiple constituencies. These can include the staff members themselves, the students, their parents, and the community, as well as the school district.*

As by far the most substantial economic cost of testing at Hillview, the allocation of staff time deserves further examination here. What does it go for?

Administrators' time was spent in a number of ways. Hillview's principal devoted some of his testing time to district-wide administrators' meetings for "in-service" on state- and district-required tests. He expended eight and three quarter hours on these sessions through the year.

More of his time on testing was given over to processing materials for these extramurally mandated measures. As described by the principal, this work included "receiving the tests, distributing them to the teachers, collecting them again, checking them over, packing them for mailing, and so on." He reported spending four and one quarter hours on these tasks in the fall and again in the spring during the conjoint administration of the Stanford Achievement Test and Otis-Lennon Intelligence Test. Similar handling of the State Assessment tests and fourth-grade proficiency test consumed three hours and an half, respectively.

But the greatest proportion of the time the principal gave to testing as spent in the review and analysis of test results. He routinely calculated year-to-year comparisons of scores for different

* On can reasonably argue that the value gained by the allocation of staff time to testing -- e.g., in more appropriate instruction; in clearer communication of students' educational status to parents, next year's teacher, and subsequent school, etc. -- is well worth the cost. Nevertheless, staff time is a cost of gaining the

classrooms and grade-levels, noted trends, and disseminated these and similar analyses to teachers. In so doing, he extended the information provided in the reports of the state or testing companies. (Note that this time is a cost of obtaining assessment information. The time the principal and teachers spent making use of test results is not included here or elsewhere in this report.) Some 42 of the principal's work hours were in test-score review and analysis through 1981-82.

A second staff member, the instructional specialist who ran Hillview's learning lab, assisted the principal in coordinating the Stanford Achievement testing. She gave 18 hours of her time to this work in the fall and once again in the spring. Her responsibilities included helping to distribute test forms; answering teachers' questions about administration procedures; assuring that all test forms were returned; and re-checking the students' answer sheets to be sure that stray pencil marks were erased, answer slots were sufficiently "bubbled in", etc.

As Table 20 shows, the principal and learning lab instructor together expended 99.75 hours on testing. For both, testing responsibilities consumed less than 5% of their school-year work time. How they allocated the time that they did spend is summarized below.

Table 25

Summary of Administrators' Annual Time
(In hours, showing % of their total time on testing)

District in-service to prepare for testing	8.75 (8.8%)
Processing test form, overseeing administration	49.00 (49%)
Reviewing and analyzing test results	42.00 (42%)
	99.75

Classroom Teachers' time on testing was spent in such diverse ways that it must be discussed more generally than that of the administrators.

As Table 24 indicates, the average (mean) time Hillview teachers spent on testing in 1981-82 was about 253 hours. Calculating annual work time as described in the footnote to Table 24, this constitutes 15.5% of a Hillview teacher's yearly work effort. Naturally, these averages mask some diversity in the allocation of time to testing. A simple listing (reveals the extent of tis variation). Below, teachers' total terms on testing per annum are displayed, together with the number of different kinds of tests that they reported giving through the year. (Here, "kind of test" refers broadly to such separate measures as a weekly spelling test, reading unit tests, reading quizzes, the Otis-Lennon, etc.) Teachers' grade levels are indicated parenthetically.

Teacher (Grade)	Number of Different Tests	Hours per Year on Testing
Fulsom (K)	8	210.5
Gardener (1)	9	215.05
Jameson (2)	10	163.91
Skoviak (2/3)	11	288.9
Fushima (3)	13	386.67
LaMarr (4)	16	250.91
Earle (4)	16	395.85
Vera (5)	19	306.05
Hurteby (5)	18	260.93
Leacock (6)	8	151.75
Coxe (6)	8	152.25

The number of different kinds of test given increases regularly until the sixth grade, where Leacock and Coxe team teach and choose to employ a variety of assignments and projects, instead of tests, for assessment. Nevertheless, in some instances, the time devoted to testing varies markedly within a grade and betwen adjacent grades. (Compare the total hours of Jameson, Skoviak, and Fushima, or of LaMarr and Earle.)

A second point worthy of note is that on the average Hillview teachers spend only about a third (34.2%) of their testing-related time in actually administering tests. Here, test administration is conceptualized to include all the classroom time from the moment when the teacher begins to give directions toward accomplishing the test until he or she moves on to the next class activity. Thus, such activities as re-arranging seating, explaining the test format, answering students' questions beforehand, distributing and picking up test papers, and so on are all included in this definition of administration time. So, too, are relaxation periods between and immediately after different portions of a test battery. (Many teachers at Hillview and elsewhere provide their children time to "cool out" or "settle down" after sections of standardized tests.) This, then, is a broad (but appropriate) operational definition of test administration. Nevertheless, the mean time devoted to these "during testing" activities in 1981-82 was about 86.5 hours of a mean total on testing.

Put another way, roughly two-thirds (65.8%) of Hillview teachers' average testing time (again, averaged across the school's eleven classroom instructors) was spent before and after classroom testing

episodes. Time before testing was, as one might expect, invested in constructing and duplicating tests, reviewing the appropriateness of questions in commercial curricular measures, reading administration directions for annual and bi-annual test batteries, and (in some instances) foregoing routine instruction to drill students on information and skills in explicit preparation for a test.* The Hillview faculty spent and average of 27.5 hours in 1981-82 (10.9% of the mean total testing time) on such "before testing" tasks.

Post testing activities -- grading, recording scores, examining and "cleaning up" special answer sheets for machine scoring, and so on -- consumed a mean time of 138.98 hours a year for the Hillview classroom staff. This constitutes 54.9% of the average of 253 testing-related hours per teacher per year.

The time that teachers devote to these before-, during-, and after-testing activities comprises by far the largest proportion of Hillview's annual testing "budget": $42,625 (or 90.5%) of the $47,085 total. Bear in mind that this is an indirect cost, one met within the routine payment of teachers' salaries.

Table 26

Summary of Classroom Teachers' Annual Testing Time

Mean time per teacher per year devoted to:		
"before testing" activities	27.5 hours	(10.9% total)
"during testing" activities	86.5 hours	(34.2% total)
"after testing" activities	138.98 hours	(54.9% total)
Mean, all testing-related activities	252.96 hours	(54.9% total)
Proportion of average annual work time testing**		15.5% total

* Instructional activities such as these were included as testing time costs only when teachers reported that they would not have conducted them were it not for the test. Routine teaching of skills covered by a test was not included in calculating staff time allocated to testing.

** See Table 24 footnotes for calculation of classroom teachers' average annual work time.

The Instructional Specialist's testing time, in her capacity as learning lab resource teacher was spent in three general ways. First, she gave placement tests in reading and math to all students new to Hillview and also elicited a writing sample from them. during the 1981-82 school year, she expended 71.3 hours on these tasks. second, in accordance with State law, she assessd the English language proficiency of incoming students when English was not the language spoken in their homes. (In some instances, the results of this assessment suggested that the writing sample and/or reading placement should be omitted.) This responsibility consumed 70 hours of her time during the year. An third, she routinely tested student sin her daily English-as-a Second-Language (ESL) class in language arts and spelling. Doing so took up 56.33 hours in 1981-82*. In all, then, the Hillview instructional specialist spent 197.63 hours on testing through the year. Using the salary rates described in Table 24, the dollar value of this time equals $2610 -- about 5.5% of Hillview's annual testing costs.

Referring once more to Table 24, it is evident that the testing efforts of the paid professional staff at Hillview were supplemented by 77.66 volunteer hours throughout 1981-82. While volunteers' time is "free", the allocation of their hours to testing constitutes an opportunity cost of Hillview's assessment program. The use of volunteer time for other tasks was forgone on behalf of testing.

For the most part, parent volunteers at Hillview helped out with standardized testing. Some asisted in proctoring; others, in the time-consuming task of examining completed answer sheets for stray marks, insufficiently darkened "bubbles" (answer markings), and

incomplete or incorrect student identification information. They also helped with such jobs as alphabetizing the forms.

Student time on testing is the last item in the overall itemization of Hillview testing costs presented in Table 24. (The rationale for including student time as a cost of testing was outlined earlier in the district-level cost accounting for Littleton.) Note that across Hillview's eleven regular classrooms, mean time per student per year is a fraction over 88 hours. This is roughly equivalent to the mean time per teacher spent in "during testing" administration (86.5 hours). But note also that on the average, nearly three hours of teacher time are required to deliver each hour of testing to the students.

Students at Hillview rarely spend cost-generating time on assessment before or after the test-taking episode. Based upon teachers' reports, the mean "before testing" time per student per year was 2.88 hours. (This of course excludes the routine teaching-learning time that precedes a test.) The mean "after testing" time per student per year was 5.34 hours. Together, these opportunity costs comprise only 9.4% of the 88.04 hours per students annual average. What is more, most of this "before" and "after" time can be traced to the two fifth grade classrooms at Hillview. Therein, students spent considerable amounts of time in explicit preparation for a State-mandated physical education assessment. From September to April, they devoted a portion of their daily physical education period to practicing exercises included on the test, exercises which would otherwise not have been part of their P.E. program. The fifth grade teachers also routinely engaged their pupils in in-class test

correction (defined here as an after-testing activity). Approximately 50% of the "before testing" and "after testing" student time investment reported school-wide occured in these two classrooms.

Finally, the general testing budget in Table 24 shows that Hillview's annual testing costs of $47,085 (all indirect costs) equal $246.52 per pupil. This may seem a large amount, but it comprises only 13.4% of Littleton District's annual per-pupil expenditure ($1836).

Table 20 and the immediately preceding discussion constitute a basic accounting of Hillview Elementary School's 1981-82 testing costs. With little additional narration, this information can be reconfigured to address a number of interesting and important questions.

Hillview's Costs for Required and Non-Required Testing

What proportion of Hillview School's yearly testing costs are incurred as a result of various testing requirements? Tables 27 and 28 provide answers to this question.

State required testing consisted of: (1) an annual State Assessment at grades 1,3, and 6; (2) the once-a-year physical performance test at grade 5; and (3) the language assessment of all potentially non-English proficient youngsters mandated in state bilingual education legislation. Collectively, these requirements feel more heavily upon the Instructional Specialists' and Principals' time, but comprised a very small proportion of the overall staff-time investment in testing. As Table 24 indicates, a mere 5% of Hillview's testing costs in 1981-82 were allocated to State-required testing.

District testing requirements are listed in Table 19 above. For Hillview, these seem at first glance to have occasioned 47% of all

TABLE 27

HILLVIEW SCHOOL - LITTLETON DISTRICT
DISTRIBUTION OF STAFF & STUDENT TESTING TIME PER YEAR
On Required and Non-Required Testing*

Each staff category cell shows:
- No. of staff members involved
- Avg. hours/staff member/year
- % Total testing time for staff by category

TYPES OF TESTING	ADMINISTRATORS' TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	VOLUNTEERS' TIME	TOTAL STAFF TIME (In Person Hours)	AVG. STUDENT TIME PER STUDENT (hours)	NUMBER OF CLASSROOMS
Required by State	1 15.75 15.8%	9 8.66 2.8%	1 70.0 35.4%		163.6 5.2%	4.46	9
Required by District	2 42.0 84.2%	11 117.66 46.5%	1 71.3 36.1%	3 24.22 93.6%	1522.2 58.2%	40.26	11
Required by School Principal		2 12.91 0.9%			25.8 0.8%	5.08	2
TOTAL REQUIRED (In person hours)	99.75 (100.0%)	1397.9 50.2%	141.3 71.5%	72.66 93.6%	1711.6 54.2%	44.46	11
NOT REQUIRED (In person hours)		1384.6 49.8%	56.33 28.5%	5.0 6.4%	1445.9 45.8%	43.57	11
TOTALS by staff category (In person hours)	99.75 (100.0%)	2782.5 (100.0%)	197.63 (100.0%)	77.66 (100.0%)	3157.5 (100.0%)		

* Required testing includes any testing mandated by someone or some agency in the organizational hierarchy above the classroom teacher.

TABLE 28

HILLVIEW SCHOOL - LITTLETON DISTRICT
DISTRIBUTION OF TESTING COSTS PER YEAR
Required & Non-Required Testing

TYPES OF TESTING	ADMINISTRATORS' TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	TOTAL DOLLAR VALUE (% Total)
Required by State	$ 253	$ 1193	$ 924	$ 2370 (5.0%)
Required by District	$ 1349	$19821	$ 942	$22112 (47.0%)
Required by School Principal		$ 384		$ 384 (0.8%)
TOTAL Required	$ 1602	$21398	$ 1866	$24866 (52.8%)
TOTAL Not Required		$21227	$ 744	$21971 (46.7%)
TOTAL by category (% Total)	$1602 (3.4%)	$42625 (90.5%)	$ 2610 (5.5%)	$46837
			District Office Testing Costs (0.52%)	+ $248
			TOTAL	$47085

1981-82 testing costs. Note, However, that among the tests required by Littleton District were various measures accompanying the reading and math text series that all teachers used. A substantial proportion of Hillview school's staff time testing costs were incurred in the use of these measures. In fact, if one excludes the time spent on them from the "required-by-District" total, that total is very nearly cut in half. Some 739 person hours are deleted from the total of 1522 spent on District-required testing, leaving about 783. This would constitute 25% of the total staff person hours devoted to testing, rather than the 48.2% shown. Instead of 52.8% of Hillview's testing costs (Table 28) being devoted to all required testing, only 31% would be.

Why consider all this? After all, the curriculum reading and math tests <u>are</u> required. While that is quite true, the issue with regard to testing requirements is usually framed in terms of testing <u>added on top of curriculum-embedded measures, on top of teachers' routine testing</u>. Teachers, for instance, sometimes argue that such testing takes up their time but provide little new information about their students. From the perspective of teachers and their advocates, then, "required testing" is often of marginal necessity. But the routine tasks associated with teaching -- such as monitoring students' learning progress, grading, and conferencing with parents -- reqire recurrent assessment. Tests intimately connected with the curriculum-in-use are a practical necessity. If some such measures were not mandated, teachers would probably need to select or devise others. In light of all this, it has been worth documenting how the required/non-required testing picture would look at Hillview were the Ginn 720 reading tests and Scott-Foresman math tests not mandated.

As matters stood, however, these tests were mandated by Littleton District. District-required testing was responsible for 47% of Hillview's 1981-82 testing costs. And slightly over half these costs resulted from mandates originating outside Hillview School.* The mean time per teacher per year devoted to required testing was about 127 hours; to non-required testing, approximately 126 hours. And notice that the typical student at Hillview spent just slightly more than half of his/her testing time, on the average, on mandated measures.

Hillview's Costs for Different Types of Testing

Tables 29 and 30 display Hillview School's 1981-82 testing costs by test type. The categories employed for typifying tests are eclectic in nature but isomorphoric with practitioners' everyday ways of talking about tests. They were identified as such in the Test Use Project's first-year exploratory fieldwork and have been employed throughout the project.

Several categories deserve brief explication. "Other, miscellaneous" testing at Hillview included: (1) the previously mentioned, State-mandated physical performance test; (2) handwriting samples requested by the principal; (3) assessment of language competence as required by State bilingual legislation; and (4) certain commercially available, diagnostic instruments employed in the early grades.

District-continuum testing consisted only of the district-developed mathematics operations tests, which seemed based on a sequence of math objectives.

Minimum competency testing took the form of a locally available "proficiency test" administered in fourth grade.**

* Two fifth-grade teachers reported that the principal-required formal penmanship samples five times a year. This was the only school-level testing mandate identified.

** The Littleton District's list of District tests indicates that proficiency testing occurs at the fourth and sixth grades. Sixth grade teachers at Hillview, however, did not report the test.

TABLE 29

HILLVIEW SCHOOL - LITTLETON DISTRICT
DISTRIBUTION OF STAFF & STUDENT TESTING TIME PER YEAR
By Type of Test

Each staff category cell shows:
* No. of staff members involved
* Avg. hours/staff member/year
* % Total testing time for staff category

TYPES OF TESTING	ADMINISTRATORS' TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	VOLUNTEERS' TIME	TOTAL STAFF TIME (In Person Hours)	AVERAGE STUDENT TIME PER STUDENT[1] (In hours)	NUMBER OF CLASSROOMS
Standardized, Norm- Referenced (Grades)	2 42 84.2%	11 34.18 13.5%		3 17.8 68.7%	513.27 16.2%	19.6	11
State Assessment Program (Grades)	1 11.25 11.3%	5 7.0 1.26%			46.25 1.5%	3.0	5
Minimum Competency (Grades)	1 4.5 4.5%	2 9.33 .67%			23.16 0.73%	3.5	2
District Continuum (Grades)		8 36.55 10.5%		1 12.33 15.9%	304.73 9.6%	6.9	8
Commercial, Curriculum- Embedded (Grades)		11 122.48 48.4%	1 71.3 36.1%	1 5.0 6.4%	1423.61 45.1%	34.5	11
Teacher Constructed (Grades)		11 55.5 21.9%	1 56.33 28.5%		666.83 21.1%	23.7	11
General Intelligence (Grades)		7 4.39 1.1%		2 3.5 9.0%	37.75 1.2%	2.9	7
Other, Miscellaneous (Grades)		5 14.37 2.6%	1 70.0 35.4%		141.83 4.5%	8.18	5
TOTALS By staff category (In person hours)	99.75 100.0%	2782.5 100.0%	197.63 100.0%	77.66 100.0%	3157.43		

[1] Note that the number of classrooms in which each type of test is administered varies, thus the proportion of time the typical student spends on each type of test varies from classroom to classroom and the average times shown cannot be appropriately added.

TABLE 30

HILLVIEW SCHOOL - LITTLETON DISTRICT
DISTRIBUTION OF TESTING COSTS PER YEAR
By Type of Testing*

TYPES OF TESTING	ADMINISTRATORS' TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	TOTAL DOLLAR VALUE (% Total)
Standardized, Norm- Referenced (Grades K-6)	$ 1349	$ 5754		$ 7103 (15.1%)
State Assessment Program (Grades 1, 3, 6)	$ 181	$ 537		$ 718 (1.5%)
Minimum Competency (Grade 4)	$ 72	$ 286		$ 358 (0.76%)
District Continuum (Grades 1-5)		$ 4476		$ 4476 (9.5%)
Commercial, Curriculum-Embedded (Grades 1-6)		$20660	$ 942	$ 21602 (45.9%)
Teacher Constructed (Grades K-6)		$ 9335	$ 744	$ 10079 (21.4%)
General Intelligence (Grades K-6)		$ 469		$ 469 (1.0%)
Other, Miscellaneous (Grades)		$ 1108	$ 924	$ 2032 (4.3%)
TOTAL by category (% Total)	$ 1602 (3.4%)	$ 42625 (90.5%)	$ 2610 (5.5%)	
			District Office Testing Costs† (0.52%)	+ $ 248
				$ 47085

* Costs of staff time are calcualted by multiplying percentage of staff time spent per category or cell (Table 29), by total dollar equivalent for staff category.

† District Office Costs pro-rated for Hillview School ($1.30 per pupil x 191 pupils = $248). These costs cannot be apportioned exactly by test type for Hillview Elementary, but see Chapter Two for a description of how Littleton District resources are allocated across different parts of the district-wide assessment program.

The "general intelligence" test category did not fall within the purview of our study of achievement testing. Teachers repeatedly mentioned it in interviews, however, and we chose to include it here to provide a more complete picture of testing at Hillview School.

With these elaborations, the findings shown in Tables 29 and 30 are self-explanatory. Notice that the largest percentage of staff and students time is devoted to tests which accompany commercial curriculum materials -- consumable test booklets linked to reading and math series, tests printed at the end of the chapter in language arts and social studies texts, etc. Considerable time was expended too, on teacher-constructed tests and quizzes (also closely tied to the curriculum), as well as on the standardized, norm-referenced Stanford Achievement Test.

Hillview's Costs for Testing in Different Subject Areas

The magnitude of Hillview School's testing costs for different subject areas is shown in Tables 31 and 32. The former reveals that Hillview educators concentrate their formal assessment efforts mainly in the basic-skills subjects. Except for administrators, all categories of participants in assessment at Hillview spend the plurality of their time on testing in math. Reading and spelling also receive larger commitments of staff and student time.

Worth noting, too, is that testing in social studies, science, and subjects categorized under "other" (such as art and music) occurs in comparatively few Hillview classrooms.* And in those where

---

* Teachers who do not test in science, social studies, art, etc. report evaluating students' progress in other ways -- through special projects, assigned reports, and routine classwork, for example.

TABLE 31

HILLVIEW SCHOOL - LITTLETON DISTRICT
DISTRIBUTION OF STAFF & STUDENT TESTING TIME
By Subject

Each staff category cell shows:
* No. of staff members involved
* Avg. hours/staff member/year
* % Total testing time for staff category

SUBJECT AREAS	ADMINISTRATORS' TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	VOLUNTEERS' TIME	TOTAL STAFF TIME (In Person Hours)	AVG. STUDENT TIME PER STUDENT (hours)	NUMBER OF CLASSROOMS Total = 30
Reading		11 52.47 20.7%	1 17.4 8.8%	1 5.0 6.4%	599.6 19.0%	12.12	11
Mathematics		11 77.11 30.5%	1 53.9 27.3%	3 15.44 59.7%	948.46 30.0%	25.11	11
Language Arts		8 24.30 7.0%	1 34.75 17.6%		229.17 7.3%	7.81	8
Spelling		8 51.42 14.8%	1 21.58 10.9%		432.97 13.7%	19.34	8
Social Studies		5 19.55 3.5%			97.75 3.1%	4.53	5
Science		5 28.0 5.0%			140.0 4.4%	5.8	5
Health - Phys. Ed		3 8.33 0.9%			25.0 0.8%	7.19	3
Other, Miscellaneous		3 8.61 1.0%	1 70.0 35.4%		95.83 3.0%	3.39	3
Multi-Subject*	2 49.87 100.0%	11 42.06 16.6%		3 8.78 33.9%	588.77 18.6%	23.93	11
TOTALS by staff category (In person hours)	99.75 100.0%	2782.5 100.0%	197.63 100.0%	77.66 100.0%	3157.55 99.9%		

* The Multi-subject category includes standardized tests which assess performance in several subject areas. Also included in this category is the general intelligence test given twice a year at the same time as (i.e., on a day contiguous with) the standardized test. Some respondents reported time devoted to the intelligence test as separate from that given to the standardized test; others did not. Thus, time devoted to both is collapsed here.

TABLE 32

HILLVIEW SCHOOL - LITTLETON DISTRICT
DISTRIBUTION OF TESTING COSTS PER YEAR
by Subject

TYPES OF TESTING	ADMINISTRATORS' TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	TOTAL DOLLAR VALUE (% Total)
Reading		$ 8823	$ 230	$ 9053 (19.2%)
Mathematics		$ 13001	$ 713	$ 13714 (29.1%)
Language Arts		$ 2984	$ 459	$ 3442 (7.3%)
Spelling		$ 2984	$ 284	$ 6592 (14.0%)
Social Studies		$ 6308		$ 1492 (3.2%)
Science		$ 1492		$ 2131 (4.5%)
Health - Phys. Ed		$ 2131		$ 384 (0.8%)
Other, Miscellaneous		$ 384	$ 924	$ 1350 (2.9%)
Multi-Subject	$ 1602	$ 426		$ 8678 (18.4%)
TOTAL by category (% Total)	$ 1602 (3.4%)	$ 42625 (90.5%)	$ 2610	$ 46837
			District Office Testing Costs (0.52%)	+ $ 248
			TOTAL	$ 47085

teachers and learners do give time to testing in these subjects, it is usually less time per year than in the basic skills.*

This concludes the itemization of Hillview Elementary School's testing costs for the 1981-82 school year. Discussion now turns to the costs of testing at Metro District's Cityside School. Once the findings of this second case study have been presented, it will be appropriate to summarize and discuss the implications of the testing-cost accountings for both schools.

Testing Costs in Metro District's Cityside School

Cityside is one of more than a hundred elementary schools in the large Metro School District. Of Cityside's 830 students, approximately 70% are Black; 28% are Hispanic; the remaining 2% is comprised of Asian, Pacific Island, and White Anglo children. Once an affluent Black neighborhood, the Cityside attendance area now ranks socioeconomically in Metro District's lowest quartile.**

Urban schools with low-income students are often portrayed as troubled environments. Cityside, however, is among the many Metro elementary schools that belie this stereotype.

Across the Cityside professional staff, the mean length of employment at the school was just under six years. Overall, the faculty averaged fourteen-and-a-half years in the field of education.

---

* This may be explained by the fact that many teachers report spending less instructional time in "the content areas" than in the basic skills. If less material is covered per year, it may not be necessary for tests to occur as frequently or to last as long.

** Metro District's socioeconomic rankings, are based upon the proportion of students families receiving Aid to Families with Dependent Children (AFDC) and the percentage of enrollment qualifying for free school lunches under federal guidelines.

A core of veteran urban teachers managed Cityside's programs, and they cited the "strong, experienced" faculty as a strength of the school. The Cityside principal concurred in this judgement. (Although new to the school in 1980-81, he had many years of leadership in other Metro District schools.)

The staff found their students capable and easy to work with. As one program coordinator put it, "we have a fairly good student body; it's not a rough school." Another with experience in schools across the District touted her Cityside position as "a plum."

The average income level of students' families qualifies Cityside for compensatory-education and other special funding under a variety of federal, state, and District categorical education programs. Chief among these are the federally sponsored Chapter I (formerly Title I) program and various supports for bilingual education. These and others provide support for additional personnel who support the work of Cityside's thirty classroom teachers. Three-hour-a-day aides (or paraprofessionals) are available for these teachers. Special program funds also support a reading resource teacher and her aide, Chapter I and Bilingual Program Coordinators, and specialists who respond to children with special learning needs.

Among the many Metro District elementary schools with compensatory education funding, cityside ranked in 1979-80 among the top 2% in reading achievement. Its sixth-grade median on the Comprehensive Tests of Basic Skills (CTBS) was then at the 56th percentile, compared to a median of the 31st percentile for all Metro District's comp. ed schools. Its scores declined to the 38th

percentile in 1980-81, but they remained above the District-wide median for schools with compensatory programs (32nd percentile, based on schools' sixth-grade medians).

The testing program at Cityside varies somewhat more from classroom to classroom than Hillview's. This occurs largely because Cityside's teachers have greater discretion over curricular testing in reading and math. Table 33 below displays the tests routinely given at Cityside Elementary.

TABLE 33

Cityside Elementary School Testing Program

Test	Grade(s)	Required by:	Administrations Per Year
Multi-Subject			
Metropolitan Achievement Test†	1 - 6	Principal	1
Comprehensive Test of Basic Skills (CTBS)	3,5	District	1
CTBS-Espanol	1,2	District	1
District Continuum Basic Skills Survey*	1 - 6	District	1
State Assessment Program	3,6	State	1
Reading			
District Reading Program†	K - 6		3 - 10
San Diego Quick Assessment†	1 - 5		1
Math			
Teacher-constructed math tests or those included in "Math for Individual Achievement" texts	1 - 6		variable
Spelling			
Teacher-constructed spelling tests; some use of commercially available word lists†	1 - 6		weekly
Language Competence			
Basic Inventory of Natural Language (BINL)	K	District	2
Moreno (Assessment of Second Language Acquisition)	K	State	1
Physical Education			
Physical Performance Test	5	State	1

† test widely administered but not in every classroom.

* The District Continuum-Based Skills Survey is required by the district at every grade. Items vary from grade to grade, covering District-defined "essential skills." The tests at grades 3 and 6 function to fulfill State requirements for minimum competency testing (and are counted as such in the following cost itemizations), although they are no different in design than those given at grades 1,2,4, and 5.

This brief description of Cityside Elementary School and its achievement testing effort provides background for the following discussion of Cityside's annual testing costs.

Cityside's Testing Costs in Overview

Table 34 provides a comprehensive look at the yearly costs of testing at Cityside Elementary. In general, the distribution of costs is quite similar to that at Hillview. The chief differences are: (1) unlike Hillview, Cityside made some direct, testing-related purchases; (2) indirect costs in administrative time were higher; and (3) costs in personnel time were distributed across a greater number of kinds of staff.

As in the Hillview overall cost accounting (Table 24), the first item in Table 34 carries district-office costs forward for Cityside's 830 pupils.

Direct dollar outlays come next in the itemization of Cityside's testing costs. At the principal's behest, Metropolitan Achievement tests were given annually. The purchase of these required $1200 per year. A basal reading series was supplemented with the Metro District's skills-oriented reading program at Cityside. it was accompanied by tests, which were consumables costing $5000 annually. The school also had a Scantron scoring machine, which automatically scored tests taken on special answer sheets. The machine was used infrequently and asystematically by individual teachers. More than the minimum number of forms were rarely purchased, an administrator reported.

Administrators/coordinators of school-wide testing at Cityside spent 375 hours in doing so during 1981-82. They performed many of the same testing-related tasks as Hillview's administrators, but

Table 34

Total Costs for All Achievement Testing in
CITYSIDE SCHOOL - METRO DISTRICT
[Enrollment = 830]

	Hours/Year (% Work Time)[2]	Dollar Equivalents[3]
District-Office Costs[1]:		
$2.64 per pupil x 830 pupils		$ 2191
Direct Costs to School:		
Purchase of Metropolitan Achievement Test		1200
Purchase of Curricular Reading Tests		5000
Purchase of Scantron Scoring Machine Forms		200
		$ 6400
Indirect Costs for School (Personnel Time):		
Administrators/Coordinators -		
Reading Resource Teacher	328.5 (19.3%)	$ 5790
Title I Program Coordinator	11.5 (0.7%)	210
Teacher Testing Coordinator	35.0 (2.1%)	472
	375.0	$ 6472
Clerical/Secretarial	10.3 (0.5%)	$ 95
Classroom Teachers -		
Average Time Per Teacher	199.2 (12.2%)	$ 2745
Number of Teachers	x 30	x30
	5975.32	$ 82,350
Instructional Specialists[4] -		
Bilingual Coordinator	156.25 (9.2%)	$ 2760
Bilingual Teacher (assists with testing)	8.08 (0.5%)	112
	164.33	2872
Instructional Aides (Paraprofessionals) -		
Aide to Reading Resource Teacher (n = 1)	109.45(20.6%)	$ 657
Aide to Instructional Specialist (n = 1)	4.58 (0.9%)	$ 27
Classroom Aides (per classroom)	39.48 (7.8%)	$ 237
Number of Classrooms	x 30	x 30
	1184.50	$ 7110
TOTAL AIDES	298.5	7794
Classroom Volunteers	92.2(??)	----
Student Time[5] -		
Average Time Per Pupil	76.1 (8.6%)	----
TOTAL COSTS FOR SCHOOL (1981-82 School Year)		$ 108,174
AVERAGE COSTS PER CLASSROOM (n = 30; avg 27.67 pupils/class)		$ 3606
COSTS PER PUPIL		$ 130.33
PROPORTION OF DISTRICT ANNUAL EXPENDITURE PER CHILD (= $1890)		6.9%

TABLE 34
Footnotes

[1] Calculations of District Office Costs are Shown in Chapter Two

[2] The "% Work Time" figures are based on respondents' report of hours worked per week before, during, and after school hours. These reported hour per week were averaged by role category across the two schools studied (Cityside and Hillview). Reported hours were within similar ranges at both schools. Work times used are as follows:

(a) For administrators, coordinators, and instructional specialists: 46 hours per week x 37 weeks per year.

(b) For clerical/secretarial personnel: 40 hours a week (roughly 22.5 work days or 180 work hours per month) x 11 months per year.

(c) For classroom teachers: 44 hours per week x 37 weeks per year = 1628 hours per year.

(d) For instructional aides: 3 hours per day per classroom x 177 school days per year = 531 hours per year per classroom.

(e) No total hours per unit or person could be ascertained for volunteers.

[3] Dollar equivalents are based upon the proportion of work time expended at the following salary estimates:

(a) For administrators and coordinators - $ 30,000 salary and fringe benefits

(b) For clerical/secretarial - $ 20, 000 salary and fringe benefits

(c) For classroom teachers and instructional specialists (except coordinators) - $ 22,500 salary and fringe benefits.

(d) For instructional aides - $ 6.00 per hour

Salaries listed under (a) are somewhat lower than the actual compensation afforded at this school, but are equivalent to estimates used in the Analysis of District Costs.

[4] Instructional specialist time reported is devoted to coordinatin and conducting achievement testing for bilingual students.

[5] Student time shown equals the time spent by the typical student in each classroom averaged across the school's regular classrooms. The percentage shown is based on 5 class hours per day (not counting the hour for lunch and recess) for 177 school days per year, which equals 885 classroom hours per school.

Cityside's greater enrollment meant that certain tasks took longer at Cityside. Furthermore, special-program funding allowed Cityside coordinators to support classroom teachers' assessment efforts in a wider range of ways.

The work of the reading resource teacher illustrates the latter point. She managed a "retrieval room" from which classroom teachers could obtain the supplementary District Reading Program materials. She ordered the tests that accompanied this program, periodically inventoried them, and conducted staff development sessions in how to use the tests and associated record-keeping forms. When class teachers needed a specific test, the reading resource teacher located it and signed it out. During 1981-82, these activities consumed 279 of the 328.5 hours that the reading resource teacher spent on testing.

Yet another of her responsibilities was to help proctor classroom testing. She spent 10 hours doing so when the District Continuum-Based Skills Survey was given and another 10 hours during CTBS testing in grades 3 and 5. Prior to the administration of the former measure, ther reading resource tescher gave a one-hour in-service session for teachers and aides which reviewed proper administration procedures.

Finally, the resource teacher saw to the purchase and distribution of the Metropolitan Achievement Test. She also answered faculty questions on how to administer and score it. These tasks required 18.5 hours of her time at the outset of the school year.

The Cityside Title I Program Coordinator assumed primary responsibility for the District Continuum-Based Skills Survey. His role consisted of obtaining the requisite test forms from the District's testing office (three hours), securing extras when a shortage appeared (fifteen minutes), "orienting" new teachers to Skills Survey

administration procedures (one hour), and planning the school-wide schedule for Skills Survey testing with the Teacher Testing Coordinator (two hours). He gave another two hours to "scheduling the set up and orientation" for teachers, and yet another half hour to arranging for supervision of half of teachers' classes while the other half was being tested.* Helping with the work of checking over students' answer sheets, alphabetizing and packaging them to be mailed for scoring took another 70 minutes of the Title I Coordinator's time, for a total of almost 10 hours on Skills Survey testing.

The Title I Coordinator also devoted an hour-and-a half annually to consulting with the Reading Resource Teacher about her orders for test materials and passing those orders on to be typed. Finally, he gave about twenty minutes to answering teachers' questions about the State Assessment measures.

A first-grade teacher at Cityside was charged with routine management of school-wide testing. This entailed the work of distributing appropriate numbers of tests and answer sheets to each teacher, collecting test materials after administration, checking over answer sheets for correct identification information, etc. She also responded to the procedural questions teachers raised in the course of testing. Altogether, the Teacher Testing Coordinator invested 35 hours in these tasks during the year of inquiry.

In all, coordination of testing consumed 375 hours of administrators' working time in 1981-82. In addition, the Reading Resource Teacher's aide assisted her with all of her testing-related responsibilities, adding an extra 109.45 hours to the staff's

---

* Metro District recommended that teachers test one-half of their class at a time, in order to assure an environment more conducive to concentration.

investment in test coordination. (See the item headed "Instructional Aides" in Table 34.) The total, 484.45 hours per year, far exceeded the time (99.75 hours) spent by Hillview administrators on coordinating and facilitating school-wide testing. On a per pupil basis, however, the difference appears less great: .58 hours per pupil at Cityside; .52 hours per pupil at Hillview. Significantly, the administrators'/coordinators' time spent at Cityside did not include an investment in extending the analyses of scores that were returned to the school. (Recall that Hillview's principal spent his time developing year-to-year comparisons for grade levels and individual classrooms.) Instead, more time was spent by the Cityside administrators and coordinators in facilitating the test-administration process. Conducting assessment in the supplementary District Reading Program, together with the more complex testing logistics in the larger school, made this necessary.

Clerical time was also a cost of testing at Cityside elementary School. Over the course of the year, a reported 10.3 hours were spent by secretarial staff in preparing the orders for the tests that the school purchased.

Teacher time at Cityside was given over to most of the same type of activities upon which teachers' testing time was spent at Hillview. And again at Cityside, there was substantial variation in the time per teacher per year allocated to testing. Seventeen of Cityside's thirty classroom teachers were interviewed during the

study.* The total time each spent on testing is displayed in Table 35 below.

Table 35
Total Time Spent on Testing by Cityside Teachers: 1981-82

Teacher (Grade)			Hours Per Year
Gonsalves	(K)		377.16
Lehrman	(K)		55.00
White	(1)		167.95
Jackson	(1)		56.38
Irvine	(1)		87.00
Prickett**	(2)	153.08	
Prickett	(1)	161.83	314.90
Moy	(2)		331.81
Hillsen	(2)		198.10
Washington**	(2)	100.46	
Washington	(3)	146.00	246.46
Benson	(3)		262.70
Krupp	(4)		299.41
Belendez	(4)		113.41
Faschinna	(5)		107.11
Ewing	(5)		248.63
Leiderman	(5)		85.91
Berriman	(6)		105.90
Smith**	(4)	155.96	
Smith	(5)	185.23	
Smith	(6)	160.40	501.59

*Although informed consent for participation in the study was gained from Metro District and Cityside School, eight Cityside teachers declined to be interviewed. Six others professed willingness to assist in the research and scheduled interviews, but their other responsibilities recurrently kept them from keeping these appointments. As a consequence, the cost accountings that follow are based upon data reported by the seventeen teachers, supplemented by estimates for those teachers who were not interviewed. In each case, the estimates were made by ascribing the mean number of hours reported by teachers at each grade level to the teachers at that grade level who were not interviewed. Further, this estimated time was divided for each non-interviewee by test type, subject matter, and mandate based on the mean proportions of time allocated to each test type, subject matter, and mandate by teachers at the non-interviewees' grade level.

**Teaches multi-grade class. Time spent on testing shown for each grade.

Teachers' annual hours on testing spanned a greater range at Cityside than at Hillview (55.00-501.59 at Cityside; 151.75-395.85 at Hillview). Moreover, the within-grade variation is much larger at Cityside. How can one account for this?

First, Cityside teachers had greater latitude in deciding how to assess student progress in reading and math. There were no required, curriculum-embedded tests in these subjects at Cityside. There were at Hillview.

Second, even though Cityside teachers used common curricular materials in reading, they tended to use those materials in different ways. According to the Reading Resource Teacher, for instance, some teachers employed the District Reading Program materials daily while others used them only once or twice a week. Greater use of the materials meant students' passed through "steps" or "levels" in the program more rapidly--and so were tested more often with program instruments.

Third, team teaching at Hillview tended to reduce the amount of within-grade variation there. In the fifth grade at Hillview, for example, one teacher did all the teaching and testing for both classes in math and science; the other, in reading and social studies. Teachers in other grades engaged in conjoint planning such that instructional schedules and rates of progress were similar. The same was not true at Cityside.

Finally, some of Cityside's within-grade variation in testing time per teacher per year is ascribable to differences in both the instructional and assessment programs for limited-English-proficient and fluent-English-proficient students. Students who spoke primarily

Spanish, for example, worked in a Spanish-language version of the District Reading Program through their early grades, and theyb were tested on a different schedule than students using the English-language version of the same program. Limited-English-proficient kindergarten children were given individually administered oral measures that fluent English-speakers were not required to take. Where the number of limited-English-proficient youngsters in a class was greater, so was the teacher time spent administering these tests.

The distribution of Cityside teachers' annual testing time was quite similar overall to that at Hillview. "After testing" activities consumed the greatest proportion of Cityside classroom teachers' time across the year (mean percentage = 53.5). But the mean proportion of time spent by Cityside teachers "during testing" (27.8%) was less than at Hillview (34.2%). And by roughly the same proportion, Cityside instructors' "before testing" time was greater (mean percentage = 18.7% as compared to 10.9% at Hillview). The classroom staff at Cityside spent more time, on the whole, preparing for classroom test administration. Several factors underly this difference.

First, Cityside teachers collectively devoted a larger proportion of their total testing time to teacher-constructed tests. Design and duplication of these measures takes time counted here in the "before testing" category.

Second, pre-administration logistics--in-service training or orientation, obtaining appropriate numbers of test forms, etc.--consumed more time at Cityside than at Hillview.

Third, more Cityside teachers reported spending time with students reviewing skills to be tested and practicing test-taking skills in advance of testing.

A summary of the main findings on the allocation of Cityside teachers' 1981-82 testing time appears in Table 36.

Table 36

Summary of Cityside Classroom Teachers' Time on Testing

Mean number of hours given to:		
"before testing" activities	37.25	(18.7% of total)
"during testing" activities	55.38	(27.8% of total)
"after testing" activities	106.57	(53.5% of total)

Total: Mean Number of Hours per Teacher per Year: 199.20
Proportion of Average Total Annual Work Time* = 12.2%
Range: 55.0-501.59 hours

Instructional Aides (or paraprofessionals') time on testing provided a substantial supplement to that of teachers' at Cityside. As Table 36 just above shows, Cityside classroom teachers allocated a mean of 199.2 hours per year to obtaining test results. This compares to a mean of about 253 hours across the Hillview faculty. But as Table 34 indicates, Cityside's classroom aides supplied (on the average) another 39.48 hours a year of staff testing time to each Cityside class. When their mean time is combined with the time of teachers, the total is an average of 238.7 hours per year of staff assessment time in each classroom.** Thus, the difference in classroom- staff testing time between Cityside and Hillview is not as great as it would initially appear.

---

*Calculation of average total annual work time is explained in a footnote to Table 34 above.

**Note, too, that Cityside students (again, on the average) receive fewer hours of testing per year than Hillview students. Using means, the ratio of staff to student hours on testing is 3.13:1 at Cityside; it is 2.87:1 at Hillview.

The time of aides is less costly that that of teachers: savings in indirect testing costs accrue from their utilization. The Cityside aides' mean time of 39.48 hours per class per year cost only $237 at aides' hourly rates. In teachers' salary, the same amount of time per class per year would have had a dollar value more than twice as high, about $546.

One might expect that a good deal of the classroom aides time was devoted to tasks before and after the test-administration episode. This was in fact the case. Altogether, Cityside classroom aides spent a mean of 26.5% (or abut 10.5 hours) of their annual time on "before testing" activities--including duplicating teacher-constructed tests, assisting in instruction explicitly undertaken for test preparation, procuring appropriate test forms for the class, etc. And, on the average they gave another 32.2% (12.7 hours per class) over the year to "after testing" tasks such as grading tests and quizzes, recording scores, returning tests to students, and checking over answer sheets prior to machine scoring. In all, then, a mean of about 58.7% of aides' testing-related time was allocated to tasks outside the test-administration episode. Still, Cityside aides, on the average, spent a substantial proportion of their time on testing in the "during" phase. (Mean for classroom aides = 16.29 hours, or about 41.3% of their mean total time.)* Their work during test administration included supervising or instructing sub-groups of students not being tested at the moment, and/or proctoring the test-taking group.

---

*Observe that on the average aides spent a higher proportion of their testing-related time in the "during" phase of testing than did teachers (mean proportion = 27.8% of teachers' mean total testing time).

They also spent time on such routine activities as distributing and collecting test booklets and answer sheets, answering students' procedural questions, and helping to re-arrange student seating at the outset and the conclusion of the administration period.*

Classroom volunteers' testing time was consumed by the same types of responsibilities often assigned to aides at Cityside. In at least two cases, volunteers shared testing tasks with both the classroom teacher and an aide.

The testing time of the instructional specialists** at Cityside Elementary School was allocated exclusively to assessment of non-English-proficient and limited-English-proficient learners. The Bilingual Coordinator conducted CTBS-Espagnol testing for students across grades three through six whose English-language competence was insufficient for them to take other school-wide, multi-subject measures. She also administered the Basic Inventory of Natural Language (BINL) throughout the year as new students who qualified for language assessment arrived at Cityside. In addition the Bilingual Coordinator taught Spanish readers in a daily class, assessing their oral and written language skills on a weekly basis. A bilingual first-grade teacher also contributed a small amount of her annual work time toward administration of the CTBS-Espagnol. In all, instructional specialists spent 164.33 hours annually on these activities.

Student time on testing averaged 76.1 hours per student per year across Cityside's thirty classrooms. Calculating annual class time at 885 hours (see Table 34 footnotes), this equals 8.6% of the yearly time available for classroom learning.

---

* Recall that by the definition in use here, these activities are all part of the test administration episode.

**The testing time of instructional specialists who taught learning disabled youngsters is omitted here as outside the domain of inquiry.

Cityside students generally spent the majority of their assessment-related hours during test-administration episodes. Mean hours per student per year in the "during" phase of testing equaled 41.78. This constituted 54.9% of the mean annual total of 76.1--substantially less than for Hillview students, where "during testing" activities consumed nearly 91% of students' average annual testing time. Conversely, Cityside students spent larger proportions of their time on testing before classroom administration began and after it was over. On the average, the typical Cityside pupil devoted 10.86 hours per year (14.3% of the mean total) getting ready to take tests and 23.48 hours yearly (30.8% of the mean total) on such "after testing" activities as in-class grading and "going over" the results of teacher-scored tests. Hillview children, in contrast, spent only 9.4% of their assessment-related time in the before-administration and after-administration phases.

Overall, Cityside's economic costs for testing in the year of the study totaled $108,174. Of this total, all but $6,400 were incurred indirectly, i.e., in the dollar values of paid staff members' time. Put another way, a little over 94% of Cityside's annual testing costs were indirect, personnel-time items.

The magnitude of the total is put in perspective by considering it on a per-pupil basis. Cityside's assessment cost per child came to $130.33 in 1981-82. The Metro School District expended $1890 per student in that school year; Cityside's per pupil testing costs come to 6.9% of this figure.

The per-pupil costs of testing at Cityside were substantially less than those at Littleton District's Hillview School ($246.52 per student). It is worth pausing a moment here to explain this difference.

Note first that Cityside's testing "expenses" were higher in several areas: District-office costs per pupil, administrators' and coordinators' time, clerical time, and direct purchases. (Hillview had no costs in the last two categories.) But in view of the entire testing "budget," these costs were only fractionally higher at Cityside.

On the other hand, Cityside teachers on the average spent less of their annual work time on testing than did Hillview teachers. And the use of paraprofessionals at Cityside (aides) resulted in savings. The factor most relevant to the per-pupil cost differential between the two schools, however, was the number of students per classroom. The number at Hillview averaged between 17 and 18 per class; the number at Cityside, from 27 to 28. Now, consider that the ratio of classroom-staff to student hours on testing was similar at both schools: 3.13:1 at Cityside; 2.87:1 at Hillview. It then becomes apparent that to provide an hour of testing to a class, the classroom-instructional staff at both schools spent roughly the same time--but that hour of testing was delivered each time to an average of about 10 more students at Cityside. It is primarily for this reason--the greater number of pupils per class--that Cityside's per-pupil annual testing costs were lower than Hillview's. Employment of aides and fewer hours of testing per pupil per year were secondary factors in Cityside's lower per pupil costs.

Table 34 and its elaboration in the preceding paragraphs have provided an overall itemization of 1981-82 testing costs for the Metro District's Cityside Elementary School. Some comparisons between Cityside's assessment costs and those in Littleton District's Hillview

School have been highlighted. The sections that follow review how Cityside's annual costs for achievement testing were distributed for mandated and discretionary testing, by test type, and by subject area.

Cityside's Costs for Required and Non-Required Testing

Table 37 itemizes Cityside's staff-time assessment costs by source of mandate. Table 38 converts these to dollar values and incorporates costs of other kinds. (Reference to Table 33 above will enable the reader to identify just which tests are required by each source.)

Here, it is simply worth underscoring that Cityside's staff-time costs for required testing were rather low, and that they were markedly lower than Hillview's. At the latter school, 54.2% of staff testing time (and 50.2% of teachers' alone) was given over to mandated testing. Even excluding Hillview's District-mandated curricular testing in reading and math, 31% of staff testing time at Hillview was invested in required measures. At Cityside, by contrast, the proportion of staff time on required assessment was a little under 15% and about 12% for classroom teachers.

The distribution of testing dollars in Table 38 reflects the staff-time allocation: the addition of Cityside's costs for testing purchases does little to change the overall picture. Some 83.3% of the annual costs of testing at Cityside were allocated to measures given at teachers' discretion.

Cityside's Costs for Different Types of Testing

Tables 39 and 40 show the distribution of Cityside's costs for testing of different types. (The test-type categorization system is identical with that used in discussing Hillview's costs, and each category is described in that discussion.)

TABLE 37

CITYSIDE SCHOOL - METRO DISTRICT
DISTRIBUTION OF STAFF & STUDENT TESTING TIME PER YEAR
On Required and Non-Required Testing*

Each staff category cell shows:
- No. of staff members involved
- Avg. hours/staff member/year
- % Total testing time for staff by category

TYPES OF TESTING	ADMINISTRATORS' TIME	CLERICAL TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	AIDES' (Para-professionals) TIME	VOLUNTEERS' TIME	TOTAL STAFF TIME (In Person Hours)	AVG. STUDENT TIME PER STUDENT (hours)	NUMBER OF CLASSROOMS
Required by State	2 3.14 7.1%		17 22.1 6.3%	1 74.0 45.0%	11 9.39 8.0%		559.20 7.1%	15.0	17
Required by District	3 19.97 22.6%		20 11.73 3.9%	2 8.2 10.0%	22 3.5 6.0%	1 5.2 5.6%	393.90 5.0%	8.6	20
Required by School Principal	3 9.83 7.8%	1 .50 4.9%	23 4.9 1.9%		23 2.61 4.6%		202.2 2.5%	2.4	23
TOTAL REQUIRED (In person hours)	95.7 25.5%	.50 4.9%	722.4 12.1%	90.33 55.0%	241.16 18.6%	5.2 5.6%	1155.30 14.6%	15.0	30
NOT REQUIRED (In person hours)	279.33 74.5%	9.8 95.1%	5252.9 87.9%	74.0 45.0%	1057.29 81.4%	87.0 94.3%	6760.32 85.4%	61.1	30
TOTALS by staff category (In person hours)	375.00 100.0%	10.3 100.0%	5975.32 100.0%	164.33 100.0%	1298.5 100.0%	92.2 100.0%	7915.6		

* Required testing includes any testing mandated by someone or some agency in the organizational hierarchy above the classroom teacher. Testing required exclusively to meet federal education program requirements has been waived for Metro District.

TABLE 38

CITYSIDE SCHOOL - METRO DISTRICT
DISTRIBUTION OF TESTING COSTS PER YEAR
Required & Non-Required Testing

TYPES OF TESTING	DIRECT DOLLAR COSTS	ADMINIS-TRATORS' TIME	CLERICAL TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	AIDES' (Para-professionals) TIME	TOTAL DOLLAR VALUE (% Total)
Required by State		$ 110		$ 5188	$ 1292	$ 624	$ 7214 (6.7%)
Required by District		$ 1036		$ 3212	$ 287	$ 468	$ 5003 (4.6%)
Required by School Principal	$ 1200	$ 505	$ 5	$ 1565		$ 358	$ 3633 (3.4%)
TOTAL Required	$ 1200	$ 1651	$ 5	$ 9965	$ 1579	$ 1450	$ 15850 (14.7%)
TOTAL Not Required	$ 5200	$ 4821	$ 90	$72385	$ 1293	$ 6344	$ 90133 (83.3%)
TOTAL by category (% Total)	$ 6400 (5.9%)	$ 6472 (6.0%)	$ 95 (0.09%)	$82350 (76.1%)	$ 2872 (2.6%)	$ 7794 (7.2%)	$105983 (2.0%)
						Plus District Office Costs	2191 (2.0%)
						TOTAL	$108174

TABLE 39

CITYSIDE SCHOOL - METRO DISTRICT
DISTRIBUTION OF STAFF & STUDENT TESTING TIME PER YEAR
By Type of Test

Each staff category cell shows:
- No. of staff members involved
- Avg. hours/staff member/year
- % Total testing time for staff category

TYPES OF TESTING	ADMINISTRATORS' TIME	CLERICAL TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	AIDES' (Paraprofessionals) TIME	VOLUNTEERS' TIME	TOTAL STAFF TIME (In Person Hours)	AVG. STUDENT TIME PER STUDENT# (hours)	NUMBER OF CLASSROOMS
Standardized, Norm- Referenced (Grades 1-6)	3 15.83 12.7 %	1 0.50 4.9 %	20 11.62 4.0 %	2 8.16 9.9 %	22* 4.49 7.6 %	2 2.6 5.6 %	400.7 5.1 %	5.54	20
State Assessment Program (Grades 3,6)	2 3.14 1.7 %		8 3.32 0.4 %		2 0.89 0.14%		34.62 0.74%	2.4	8
Minimum Competency (Grades 3,6)			8 6.10 0.8%		2 5.98 0.9 %		60.79 0.76%	5.5	8
District Continuum (Grades 1,2,4,5)	3 13.97 11.2 %		20 5.76 1.9 %		9† 4.29 3.0 %		195.71 2.5 %	4.7	20
Commercial, Curriculum-Embedded (Grades K-6)	2 139.67 74.4 %	1 9.8 95.1 %	30 69.80 35.0 %		31§ 18.25 43.6 %	3 26.95 87.7 %	3029.89 38.3 %	21.7	30
Teacher Constructed (Grades 1-6)			26 119.9 52.2%	74.0 45.0 %	26 18.8 37.7 %	1 6.16 6.7 %	3685.33 46.5 %	48.1	26
Other, Miscellaneous (Grades K-6)			20 17.14 5.7 %	2 37.0 45.0 %	9 10.28 7.1 %		509.22 6.4 %	10.3	20
TOTALS By staff category (In person hours)	375.0 100.0 %	10.3 100.0 %	5975.32 100.0 %	164.33 99.9 %	1298.5 100.0 %	92.22 100.0 %	7915.7		

* Aide time includes 13 hours spent annually by aide to Reading resource Teacher in coordinating and proctoring, and 4.58 hours spent in similar duties by an aide to a bilingual specialist. Omitting these times, aides in 20 classrooms spend an average of 3.5 hours on standardized, norm-referenced testing.

† Aide time includes 1- hours spent annually by aide to Reading Resource Teacher in proctoring test administration. Omitting this time, aides in eight classrooms spend an average of 3.6 hours annually on testing associated with district continuum testing.

§ Aide time includes 81.45 hours spent annually by aide to Reading Resource Teacher in distributing, organizing, inventorying and re-ordering reading test materials. Excluding this time, aides in 30 classrooms spend an average of 16.1 hours annually on testing that is embedded with commercially available curriculum materials.

Note that the number of classrooms in which each type of test is administered varies; thus, the proportion of time the typical student spends on each type of test caries ffrom classroom to classroom and the average times shown cannot be appropriately added.

TABLE 40

CITYSIDE SCHOOL - METRO DISTRICT
DISTRIBUTION OF TESTING COSTS PER YEAR
By Type of Testing*

TYPES OF TESTING	DIRECT DOLLAR COSTS	ADMINIS-TRATORS' TIME	CLERICAL TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	AIDES' (Para-professionals) TIME	TOTAL DOLLAR VALUE (% Total)
Standardized, Norm- Referenced (Grades 1-6)	$ 1200	$ 822	$ 5	$ 3294	$ 287	$ 592	$ 6200 (5.7%)
State Assessment Program (Grades 3,6)		$ 110		$ 329		$ 11	$ 450 (0.4%)
Minimum Competency (Grades 3,6)				$ 659		$ 71	$ 730 (0.7%)
District Continuum (Grades 1,2,4,5)		$ 725		$ 1565		$ 234	$ 2524 (2.3%)
Commercial, Curriculum-Embedded (Grades K-6)	$ 5000	$ 4815	$ 90	$28822		$ 3398	$42125 (38.9%)
Teacher Constructed (Grades 1-6)				$42987	$ 1292.50	$ 2935	$47214.5 (43.6%)
Other, Miscellaneous (Grades K-6)	$ 200			$ 4694	$ 1292.50	$ 553	$6739.50 (6.2%)
TOTAL by category	$ 6400	$ 6472	$ 95	$82350	$ 2872	$ 7794	$105,983
(% Total)	(5.9%)	(6.0%)	(0.09%)	(76.1%)	(2.6%)	(7.2%)	
						District-Office Costst (2%)	$ 2191 (2.0%)
							$108174

* Costs of staff time are clacualted by multiplying percentage of staff time spent per category or cell (Table ???), by total dollar equivalent for staff category.

† District Office Costs pro-rated for Cityside School ($2.64 per pupil x 830 pupils = $2191). These costs cannot be apportioned exactly by test type for Cityside Elementary, but see Chapter 39 for a description of how Metro District resources are allocated across different parts of the district-wide assessment program.

One note of explanation is necessary. Recall that the same series of tests (the District Continuum-Based Skills Survey) falls under two categories in these tables. At grades 3 and 6 the Skills Survey functioned to meet state requirements for minimum competency testing. At grades 1, 2, 4, and 5, tests in the Skills Survey is counted as a District Continuum test. (At all grades, the Skills Survey assessed students' learning of skills on District reading, math, and language arts continua that have been designated as "essential".)

Overall, Cityside staff gave the largest proportion of their assessment time (46.5%) to teacher-constructed measures. Over half of classroom teachers' time on testing occurred in conjunction with these. Another 38.3% of the staff's time allocation to testing took place in the context of commercial, curriculum-embedded measures. (The plurality of aides' time was spent on these.) Note too, that the average time spent on testing per student per year was also highest for these twp types of measures.

As Table 40 indicates, 82.5% of Cityside's direct and indirect costs were incurred for these teacher-constructed and commercial, curriculum-embedded testing. This was higher than at Hillview, where commercial and teacher-made curricular measures still consumed a substantial 67.3% of the annual resources given to testing. (As reference Table 29 shows, the Hillview staff time commitment was larger for commercial curricular testing, lower for teacher-constructed tests--just the reverse of Cityside's.)

Cityside's Costs for Testing in Different Subject Areas

The distribution of Cityside's staff-time on assessmenbt in different subjects is displayed in Table 41. Table 42 converts these to dollar values and adds direct-purchase testing costs.

TABLE 41

CITYSIDE SCHOOL - METRO DISTRICT
DISTRIBUTION OF STAFF & STUDENT TESTING TIME
By Subject

Each staff category cell shows:
- No. of staff members involved
- Avg. hours/staff member/year
- % Total testing time for staff category

SUBJECT AREAS	ADMINIS-TRATORS' TIME	CLERICAL TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	AIDES' (Para-professionals) TIME	VOLUNTEERS' TIME	TOTAL STAFF TIME (In Person Hours)	AVG. STUDENT TIME PER STUDENT (hours)	NUMBER OF CLASSROOMS Total = 30
Reading	2 139.66 74.5%	1 10.3 100.0%	28 54.61 25.6%	1 74.0 45.0%	26 15.31 30.7%	1 11.6 12.6%	2302.42 28.8%	9.43	28
Mathematics			27 67.58 30.5		25 15.51 29.9%	2 33.06 71.8%	2278.38 28.6%	21.01	27
Language Arts			16 25.42 6.8%		10 3.63 2.8%		443.0 5.5%	18.71	16
Spelling			22 54.25 20.0%		18 11.17 15.5%	1 9.17 10.0%	1403.67 17.6%	25.83	22
Social Studies			10 17.65 2.9%		6 4.12 1.9%		201.20 2.6%	10.33	10
Science			5 16.4 1.4%		2 0.63 0.09%		83.25 1.0%	4.33	5
Health - Phys. Ed			6 16.55 1.7%		6 9.52 4.4%		156.47 2.0%	30.28	6
Other, Miscellaneous			6 40.27 4.0%	1 74.0 45.0%	4 10.34 3.2%		356.96 4.5%	0.39	6
Multi-Subject	3 31.90 25.5%		26 16.24 7.1%	2 8.16 10.0%	28 5.39 11.6%	2 2.6 5.6%	690.45 9.4%	9.62	26
TOTALS By staff category (In person hours)	375.0 100.0%	10.3 100.0%	5975.32 100.0%	164.33 100.0%	1298.5 100.09%	92.22 100.0%	7915.8		

TABLE 42

CITYSIDE SCHOOL - METRO DISTRICT
DISTRIBUTION OF TESTING COSTS PER YEAR
by Subject

TYPES OF TESTING	DIRECT DOLLAR COSTS	ADMINIS-TRATORS' TIME	CLERICAL TIME	CLASSROOM TEACHERS' TIME	INSTRUCTIONAL SPECIALISTS' TIME	AIDES' (Para-professionals) TIME	TOTAL DOLLAR VALUE (% Total)
Reading	$ 5000	$ 4822	$ 95	$ 21081	$ 1292.50	$ 2393	$34683.50 (32.1%)
Mathematics				$ 25117		$ 2330	$27447. (25.4%)
Language Arts				$ 5600		$ 218	$ 5818 (5.4%)
Spelling				$ 16470		$ 1208	$17678 (16.3%)
Social Studies				$ 2388		$ 148	$ 2536 (2.3%)
Science				$ 1153		$ 7	$ 1160 (1.1%)
Health - Phys. Ed				$ 1400		$ 343	$ 1743 (1.6%)
Other, Miscellaneous	$ 200†			$ 3294	$ 1292.50	$ 249	$ 5035.50 (4.6%)
Multi-Subject	$ 1200	$ 2749		$ 5847	$ 287	$ 904	$ 9888 (9.1%)
TOTAL by category	$ 6400	$ 6472	$ 95	$ 82350	$ 2872	$ 7800*	105989
(% Total)	(5.9%)	(6.0%)	(0.09%)	(76.1%)	(2.6%)	(7.2%)	
						Plus District-Ofice costs	$ 2191 (2.0%)
							108180

† Expenses for scantron scoring forms are ascribed to "other miscellaneous" category

* Total is slightly larger for this category than in previous tables as a result of rounding off percentages in Table 37 (Dollar amounts here are based upon those time allocation percentages.)

As at Hillview, Cityside's staff-time testing costs were concentrated in the basic skills subjects of reading, math, and spelling. Also, as at Hillview, the basic skill of language arts (grammar, writing, oral communication--but excluding spelling here) received a substantially lower proportion of the Cityside staff's total testing-time investment than the other basic skills.* Another similarity between the two schools--a corollary to the basic-skills testing emphasis--was evident in the comparatively low allocation of Cityside staff time to testing in the areas of science and social studies.

It is also worth noting that Cityside's staff-time commitment in multi-subject testing was about half Hillview's (9.4% as compared to 18.6% of total annual staff assessment time).**

Through the three sections immediately above, the intent of discussion has been to highlight general patterns in the distribution of Cityside Elementary School's annual testing costs and to compare salient patterns of resource allocation to those found at Hillview School. At this point, reporting turns to a summary and discussion of principal findings.

Summary

Formal interviews and supplemental fieldwork at two elementary schools provided a comprehensive picture of their annual costs for achievement testing. Findings of principal interest are highlighted here.

*Many teachers interviewed at both schools expressed a preference for non-test assessment strategies in language arts, but interviewers were asked to include regular, formal writing assignments among language arts testing.

**Multi-subject tests at Cityside included the Metropolitan Achievement Test, the Comprehensive Tests of Basic Skills, the District Skills Survey, and State Assessment measures. The last two of these cover exclusively basic-skills subjects; the first two concentrate heavily upon them.

Overall Costs

- At a large, urban elementary school (Cityside) serving a low-income enrollment of 830, annual costs for achievement testing of all types in all subjects were $108,174, or $130.33 per pupil.

- At a small, suburban elementary school (Hillview) serving a relatively high-income enrollment of 191, annual costs for achievement testing of all types in all subjects were $47,085, or $246.52 per pupil.

- Nearly all of these costs were incurred indirectly as a result of staff time spent on testing.

- The single largest item in each school's annual testing "budget" was the time that classroom teachers gave to assessment, an indirect cost of testing borne by the school districts.
  (Teacher time on assessment as a proportion of total annual testing costs: Hillview=90.5%; Cityside=76.1%.)

Staff Time

- Total administrator/coordinator time per year on testing:
  Hillview = 99.75 hours/year
  .52 hours/year/pupil
  Cityside = 375 hours/year
  .58 hours/year/pupil

- Mean annual time per teacher per year on testing:
  Hillview = 252.98 hours (15.5% annual mean work time)
  Cityside = 199.2 hours (12.2% annual mean work time)

- Paid para-professional (aide) time per classroom per year:
  Hillview = none present
  Cityside = 39.48 hours

- Volunteered time (both schools) and clerical time (Cityside) were incidental in magnitude.

- Classroom teachers at both schools spent more than two-thirds of their testing-related time prior to and after the classroom testing episode.

Distribution of Teacher Time

- Proportion of total teacher time per year on testing required by supraordinate individuals and agencies:
  Hillview = 50.2%
  Cityside = 12.1%

- Types of testing consuming greatest proportions of teachers' testing time:

	Hillview	Cityside
Teacher-constructed	21.9%	48.4%
Commercial curriculum	45.1%	35.0%
Norm-referenced, standardized batteries	13.5%	4.0%

- School subjects receiving largest proportions of teachers' annual testing time:

	Hillview	Cityside
Reading	20.7%	25.6%
Math	30.5%	30.5%
Spelling	14.8%	20.0%
Multi-subject test batteries	16.6%	7.1%

Student Time

- Average time per student per year spent on all achievement testing in all subjects (and percent total annual classroom instructional time of 885 hours):

  Hillview = 88.04 (9.9%)
  Cityside = 76.10 (8.60%)

- Average student time per student per year on testing required by individuals and agencies supraordinate to the classroom teacher (and percent of mean total):
  Hillview = 44.46 hours (50.5%)
  Cityside = 15.0 hours (19.7%)

- Average student time per testing per year on subjects in which typical student spends most testing time (shown in hours per year):

	Hillview	Cityside
Reading	12.12	9.43
Math	25.11	21.01
Spelling	19.34	25.83
Multi-subject test batteries	23.93	9.62

Discussion

Heretofor, very little has been known about the level of schools' economic investment in the achievement testing process. The findings reported in this section, therefore, merit attention simply for their descriptive value. They provide a first, comprehensive look at the magnitude of elementary schools' testing costs. And they yield a detailed portrait of how much time teachers and students spend on testing of different types.

These findings become more useful, however, when one has some sense of whether the magnitude and distribution of these particular two schools' testing costs are typical or unique. Results of the Test Use Project's 1981 national survey allow this issue to be addressed in a general way.

Survey questionnaires went to teachers in a nationally representative sample of districts and schools across the United States. Those in the upper elementary grades were asked to "compile a complete list of tests given to assess or evaluate your students" in reading and math. Teachers were directed to report the number of times per year a "typical student" took each test listed and the "approximate time for (the) typical student to complete one." Responses to these questions, then, offer a national view of students' annual testing time in reading and math.

Table 43 summarizes the survey data in juxtaposition to the findings for Cityside and Hillview Elementary Schools. Therein, it is seems that Cityside students are a fraction below the national average for reading testing. Otherwise, Hillview and Cityside (at least in

TABLE 43

Average Hours Per Student Per Year Spent in Reading and Math Testing:

Comparison of Hillview and Cityside to National Survey Data

	Nation-Wide	Hillview	Cityside
Reading	9.93	12.12	9.43
Math	12.47	25.11	21.01
Total	22.40	37.23	30.44

math) appear to be "high testing" schools. Of course, teachers in the national survey sample were not asked to report student testing-related time spent before or after test administration. They were only directed to report on test-taking time, How would the national averages look if they were "adjusted" to incorporate an estimate of student time spent before and after testing? And how would student testing time in the two case-study schools compare?

Table 44 answers these questions. In that table, the survey averages for hours per student per year in reading and math testing have been adjusted upward. The adjustment was made by averaging the proportions of their meaning testing time students at Hillview and Cityside spent during test administration (91% at Hillview; 55% at Cityside, for an average of 73%). Then the mean times reported in the survey were considered as 73% of the total time actually spent on

TABLE 48

ADJUSTED COMPARISON*

Average Hours Per Student Per Year Spent
In Reading and Math Testing

	Nation-Wide	Hillview	Cityside
Reading	13.6	12.12	9.43
Math	17.08	25.11	21.01
Total	30.68	37.23	30.44

*See text for a description of the adjustment process.

testing, and an appropriate amount of time for before-administration and after-administration testing-related activities was added. With this "guesstimate" adjustment, Hillview and Cityside students appear to spend a bit less than the national average time on reading testing but a bit more than the average on math testing. Cityside's total is quite near the adjusted national average; Hillview's, seven hours higher.

Although this comparison is admittedly a rather crude one, it does at least hint that the amount of testing at the two case-study schools (especially in the basic skills) probably does not diverge dramatically from the amount of testing conducted in many other elementary schools in the nation.

Further support for this cautious claim can be found in survey findings on the allocation of student testing time by test type. The survey showed that in the upper-elementary grades, the greatest proportions of students' annual testing time were devoted to school- or teacher developed measures (35% - 37%). These figures are consonant with the findings in the two case-study schools (Compare Table 1, on page 4 in the Introduction to Tables 29 and 39 earlier in this chapter).

This discussion is certainly not an attempt to argue for the generalizability of the findings reported in this chapter. It is merely to put them in perspective. And the perspective suggested here is this: until further research indicates otherwise, it is appropriate to view the levels and costs of testing reported here as not atypical, as probably "in the same ballpark" with the levels and costs of testing in a good many other American elementary schools.

But what can one conclude from the findings from Hillview and Cityside Elementary Schools?

First, these findings suggest that testing does not impose an especially great burden on students' instructional time. Students in the two case-study schools spent about 9%-10% of their annual classroom instructional time on testing of all types in all subject areas. (This comes to an average of two or two-and-a-half hours per week.) Furthermore, some 60%-70% of this time was spent on testing closely linked (in intent at least) with content and process of teaching-learning, i.e., with teacher-constructed and commercial curricular testing. Assuming that regular assessment is an important part of good teaching, the scope of student time on testing certainly seems within a reasonable range.

Nor do the costs of assessment in teacher time seem especially great. Assuming that a typical elementary teacher spends 44 hours a week on job-related activities over 37 weeks a year (as teachers in the two case study schools reported doing), teachers seem to spend on the average of about 12%-15% of their yearly work time on testing. This amounts to some five-to-seven hours a week, a good bit of it spent outside of school hours on grading tests and recording test scores. This is not an inconsiderable amount of time. But it seems important to note that much of this time was invested in curricular testing (about 87% at Cityside; about 70% at Hillview). And this testing was undertaken either at teachers' discretion or with their consent (in the case of Hillview's commercial curricular measures in reading and math). Testing divorced from the curriculum and required by teachers' supraordinates consumed about 15%-30% of their total testing time -- or about 2% of their work time at Cityside and 5% at Hillview. As the next chapter will indicate, many teachers report frustrations and aggravations in conjunction with such non-curricular, required types of assessment as annual or biannual standardized testing and State Assessment. They may entail subjective costs for teachers disproportionate to the amount of teachers' time they consume. This is certainly an important consideration. But in a literal, objective sense, the time-costs of testing which is both required and divorced from routine teaching-learning are not large.

Third, it deserves reiterating that the direct costs of testing do not appear to be great. Even if districts and schools were to cut back sharply on the amount of testing they conduct, they would not find themselves with a vast sum of re-allocatable dollars. A far

greater proportion of districts' and schools' "expenses" for testing are incurred indirectly through the time staff members devote to assessment.

Fourth, elaborating on a point made earlier, the elimination of mandated testing would probably save only very modest amounts of school-level educators' time. State-mandated testing at the two schools studied consumed only 5.2% (at Hillview) and 7.3% (at Cityside) of the total yearly staff hours devoted to testing, hours which themselves constituted a small proportion of staff members' work time across the school year. District requirements comprised only another 5.6% (at Cityside) and 25% (at Hillview, excluding curricular testing requirements) of this already small proportion.

Two key issues of relevance for educational policy are suggested by the data presented here.

Districts (and perhaps schools) should consider ways of making curricular testing more efficient. The greatest cost districts and schools appear to bear for testing is the opportunity cost of teacher time. Teachers, in turn, spend the greatest proportion of their time in curricular testing. Districts and schools interested in enconomizing on assessment, therefore, should probably focus on finding ways to reduce the time teachers spend in constructing their own tests and in scoring these and other curricular measures. Item-banking and the use of computer scoring and computer analysis of test scores should be considered. These and similar procedures may have larger initial costs, but over the years they could free substantial proportions of teacher time for classroom instruction.

More broadly, the issue of test quality emerges as central in these findings. The questions "How much testing is going on?" and "What does it cost?" seem to be less important, in light of the findings presented here, than the question "How good are the tests being used?" Teachers spend substantial proportions of their assessment time on teacher-constructed and commercial, curriculum-embedded measures. Teachers also report considering these tests heavily in making instructional decisions. (Refer to Tables 2 and 3 in the Introduction to this report, which show survey findings in support of this point.) Yet, we know very little about the quality of these types of tests. We do know, however, that most teachers receive little pre- or in-service training in test construction or test selection. (On the Test Use Project's national survey, 80% of the teachers responding indicated that they received no staff development in these areas. Other CSE work suggests that teachers receive little pre-service training in assessment.) While the costs of testing seem modest or small, the impact of curricular test results certainly is not. The quality of curricular testing, then, merits further attention.

PSYCHOLOGICAL COSTS: TEACHER ATTITUDES TOWARD TESTING

Toward the close of each interview on staff members' testing time, the CSE researcher asked a series of specific questions about potential concerns and anxieties associated with testing. The questions were sturctured to discern whether these anxieties were borne by teachers, students, administrators, or others. Relevant commentary offered by teachers and administrators during early stages of the interview was also recorded directly on the interview form. These responses have been analyzed and categorized in terms of their dominant thematic content. The findings indicate that--at least in the schools--testing and the use of test results do not cause deep worry or distress; some aggravation, rather than anxiety, appears to be the principal psychological cost of testing. The nature of this aggravation is reflected in teacher concerns about test utility, appropriateness of tests and their uses, testing effects, and impact on instructional time. Each of these concerns is elaborated below.

Test Utility

Virtually every teacher interviewed at elementary Cityside commented, explicity or implicity, on the utility of some of the tests in use at their school. Fourteen teachers made very explicit comments on this topic, which suggests that having to administer tests of little direct use to teachers is a widespread concern at Cityside. Many of the negative comments reflected problems with tests that teachers are required to administer, usually norm-referenced or minimum competency tests, or tests associated with the reporting

tests that teachers are required to administer, usually norm-referenced or minimum competency tests, or tests associated with the reporting requirements of externally funded programs. These comments cut across all grade levels at Cityside. They range from simple statements asserting a general lack of test relevance to comments suggesting differential value of specific parts of a specific testing program.

In contrast to Cityside, teachers at Hillview made few direct comments about test utility. In fact, only two teachers at Hillview, mentioned such a concern. The concerns about test utility expressed by Cityside teachers, categorized by theme, are detailed below.

Lateness of test score reports; Cityside: Five educators at Cityside commented on the lateness or non-receipt of test results. Of the test required for assessing limited-English-proficient (LEP) students' language dominance, the Bilingual Coordinator noted:

> (it has) rather dubious value. There is a delay in getting the scoring back. You wait four to six weeks to get a return (and) by the time you get the results back, you've forgotten the individual child.

Similarly, one of the first-grade teachers noted that she never sees the results of the Comprehensive Tests of Basic Skills (CTBS) Espa ol, which is required for students in the school's bilingual classes, "nor are they ever given to the students or their teachers in the next grade." This teacher generally felt that she has to give a lot of tests but "gets nothing back." One of the grade two teachers at Cityside commented that she can get the CTBS Espanol results if she asks for them, but "the results come back too late" to have any instructional use. The bilingual coordinator also emphasized this

problem in her comment that "the kind of test we give at the end of the school year (e.g., CTBS Espanol), the teachers never see the results." The third-grade teacher preferred her own tests over more formal measures because of their immediate feedback potential.

Discussing the Continuum-Based Skills Survey (CBSS) which is administered across all grades at Cityside, one of the fifth-grade teachers noted that:

> the results come back too late. I don't know who they will benefit. (I) can't wait (for the scores) to do (student) grouping. I don't really use the test scores.

Lack of relevance or test redundancy; Cityside: Six teachers at Cityside commented on the problem of test relevance or actual redundancy. For example, one of the first-grade teachers noted that the school-required Metropolitan Achievement Test (MAT) does not help her with the kinds of instructional or classroom management decisions she has to make early in the school year, although it may later "back up what I've (already) done" in terms of decisions about student diagnosis and grouping in reading and math on the basis of less formal measures. One other colleague in the first grade amplified this issue by asserting that there are too many tests that "basically tell me the same thing."

Concerns at Cityside with lack of test relevance appeared to be a problem for some of the upper grade teachers as well. Discussing the MAT, a fourth-grade teacher observed that she used this test because she:

> didn't have a choice. (I) didn't find it helpful. It was a good idea to have an achievement test, but (on) this one (the student scores were) so low. They (the students) function so much better than (the scores would indicate).

Also commenting on the MAT, a fifth-grade teacher noted that the "results aren't worth the time it takes," and went on to describe the results of the CBSS and CTBS in similar terms. According to this teacher:

> One year-end test is enough. (We) need one formalized test that is useful. Two tests (are) redundant and take time away from the program.

This concern was shared by a second fifth-grade teacher, who felt that the MAT "took too much time and I didn't agree with the results." A sixth-grade teacher also observed that the MAT "was a waste (and) I didn't agree with the results." Two teachers at Littleton District's Hillview School who chose to comment on test utility offered similar remarks regarding certain tests that they were required to administer.

Differential value of parts of a testing program; Cityside:

Five teachers at Cityside made reference to the value of tests associated with the Developmental Reading Program (DRP), which is used by many teachers in the school. All of these comments indicated that the teachers in question saw no value in administering unit pretests. Most of these teachers simply admitted that they use only the unit posttests. One of the first-grade teachers went on to justify this practice:

> I don't waste time on the pretest...I only give them the posttest (and) if they pass I move them on to the next step. If they don't pass, they go over the things they miss,...then go on to the next step. It's great for diagnostics.

It would be inacurate to say that the pre-tests associated with the DRP create a psychological cost for teachers at this school: teachers can simply omit them. However, that several regularly do so suggests that dollars invested in pre-tests may not be a wise investment for all teachers.

Appropriateness of Tests and Their Uses

As was the case with test utility, virtually every teacher interviewed at Cityside had something to say about the appropriateness of tests and/or the the uses to which they are put. About a dozen of these teachers, covering most grade levels, made very explicit statements reflecting concerns about test/test use appropriateness. Teacher commentary in this category, while a great deal of it was negative, also tended to show that teachers at Cityside are not bothered by all forms of testing. Nor do Cityside teachers tend, to single out tests as inappropriate on the basis of their generic features (e.g., norm-versus criterion-referenced).

With the teachers in Hillview, a different kind of picture emerged. Here only about half of the eleven teachers commented directly on the appropriateness issue. And in each case the comment reflected a concern about manner in which a test score was used and the effect of its use on students and teachers.

Most of the Cityside comments on appropriateness fell into the following categories.

Ease/difficulty of tests; Cityside: Seven teachers at Cityside made statements about the ease or difficulty of a test or kind of test. In terms of minimum competency testing, for instance, the

School's Bilingual Coordinator noted that there is a "need for a test like the CBSS, (though) it should be more of a challenge (for the students)." One of the first grade teachers amplified this attitude toward minimum competency testing as follows:

> The CBSS, I think, should be harder...I wouldn't eliminate the CBSS, but I'd revamp (it) to where, instead of having minimal (skills), it would have maximum (competencies).

Three of the second-grade teachers agreed. One commented that the "CBSS (is) not useful. There is no worthwhile feedback." For another the Skills Survey "is too easy, not valuable," while the third felt that "the Survey could be better...it doesn't tell me how far the student can go."

Similar comments were made about some of the norm-referenced tests administered at Cityside. The Bilingual Coordinator observed that the "CTBS Espanol is far more difficult (than the CBSS), which is very minimal." This specialist was very concerned about the disparity of difficulty levels between the two tests.

The first-grade teacher quoted above believed that tests <u>like</u> the Skills Survey and CTBS (i.e., minimum competency and norm-referenced) served justifiable purposes, but felt that the purposes were not adequately fulfilled by these two particular tests. Discussing the CTBS, which was once (but no longer) required on a school-wide basis, this teacher commented:

> That's one thing the CTBS had that was good; it went far beyond what (the students) should know. But I didn't like the CTBS because it didn't start at a low enough level; it was too hard.
> So you need (a test) that starts at very minimal level and goes up beyond what (students') capabilities are, so you really get a true picture of what the potential is of the best and of the slowest.

A second-grade teacher similarly criticized the CTBS and the Skills Survey. The CTBS, she opined, is:

> too hard for most (students). They are frustrated. The Skills Survey is silly. It is costly and doesn't give a true picture.

One of the fourth-grade teachers agreed in stronger terms:

> The Skills Survey is not timed. All but three students finished. One girl got them all wrong. All she did was mark it; she wasn't even trying. It's the same when we give the CTBS. (A certain student) got the highest score, and he couldn't read. He is now in EH. I know he can't do it. He guessed.

This kind of problem was also recognized by the school principal, who is concerned about the CBSS because it has "no norming data (and has) low-level expectancy." Further, because the CTBS is no longer required school-wide, and because the principal sees some value in generating school-wide norm-referenced data, "that's why I spend $1200.00 for the MAT."

While some teachers at Cityside do see a need for minimum competency and group-administered, norm-referenced tests, they are not particularly pleased with the tests being used for these purposes. Comments amplifying their frustration appear below.

Technical problems; Cityside: Three teachers and the Title I Coordinator commented on this issue. One of the second-grade teachers criticized the CTBS Espanol because "some of the words don't translate into Spanish...(and) the print is too small...(the test) is not testing Spanish skills." A fifth-grade teacher noted similar problems with the English-language version of this test:

> (The test) vocabulary is a problem for (the students). Some of the explanations are (written in language) for adults. The test is a contradiction. (It) makes criminals of us all. It's unrealistic. It makes us all cheat.

Discussing another kind of technical problem, that of score reporting format, this same fifth-grade teacher observed that:

> There has to be a better way of reporting the scores to the teachers so they can be used...I would like to get a print out on a sheet at the beginning of the year which show all the Skills Survey and CTBS results...so I can see it all together at a glance. To have to go to everyone's cumulative file is very tedious;...someone in the school, whether coordinator, principal, or whoever is in charge, should get it all together.

That no one in Cityside, "gets it all together" was corroborated by the vice principal. Describing what was a frustrating experience for him as an administrator and for his teachers as well, he commented that:

> Some teachers want to know how students did, because the printouts aren't going to come back until school is out. If they want to know, we have a hand-scoring key if they want to do this. No one interprets school-wide.

The fifth-grade teacher who cited the concern with CTBS noted above pointed out another problem with some of the tests administered at Cityside. Teachers are very concerned because they need much more information on what the various tests mean, their "validity and correlation with other tests." Another fifth-grade teacher commented that "testing is not as controlled as it was twenty-five years ago. We would have inservice to make sure you knew what you were doing." This was also a concern for the vice principal, who commented that teachers at Cityside, in general, need more explanation from the Metro District's research and evaluation office about what the various test scores mean.

Tests viewed favorably; Cityside: Four teachers at Cityside spoke of the kinds of tests that are viewed more favorably. The

Bilingual Coordinator, for instance, discussing a Spanish reading test she developed herself, noted that this kind of testing

> is not time-consuming. It is something I can get feedback on immediately. It isn't disruptive; it's a very satisfactory, necessary instrument.

In terms of diagnostic information on students' reading ability, one of the first-grade teachers described the diagnostic value of the San Diego Quick Assessment as follows:

> I give the San Diego (Quick Assessment), which takes about thirty seconds per child (and) it's pretty accurate...one of the most accurate I've ever seen. It's something I do at the beginning of the year. You can do the whole class in fifteen or twenty minutes.

One of her colleagues strongly agreed. "I don't mind giving (the San Diego) because it doesn't take much time and it's useful." A fifth-grade teacher concurred that the San Diego Quick Assessment "is useful when you want to place a new student."

Recall also that many teachers at Cityside viewed the unit posttests of the Developmental Reading Program positively, and that some teachers also saw the value of the information they felt they could obtain from a good minimum competency test or a good norm-referenced test, though they were concerned about problems with the two tests actively used in the school for these purposes...the CBSS and the CTBS.

Effects of Testing

Most of the teachers at Cityside commented on problems arising from the effects of testing on students or teachers. And at Hillview, nine of the eleven teachers interviewed spoke about the effect that testing has in fostering student anxiety. Half the Hillview

interviewees also expressed concerns with pressures that testing can generate for teachers.

Student anxiety; Cityside and Hillview: A majority of the teachers at Cityside were concerned about tests causing students either to become very wound up and/or to become tired and enervated. In this regard, some of the teachers described efforts to incorporate student "wind-down" time after a testing period by scheduling the test immediately before recess. When this was not possible, they said they generally gave their classes about fifteen minutes (taken out of instructional time) to relax and get over the effects of testing.

About a half-dozen teachers at Cityside cited testing as a generally frustrating experience for their students. One first-grade teacher specifically refered to the MAT as "too tiring and frustrating," a view for which she found evidence in students "breaking their pencils" to try to avoid taking a test. One of the third-grade teachers mentioned that her "third-grade students get too many tests, often several at about the same time." This teacher saw her students becoming restless as the Spring testing period wore on; "testing time and its effects take a long time to wear off," she said.

One of the second-grade teachers described certain kinds of tests and their effects on her students as follows:

> The ongoing tests like the District Reading Program...aren't identified as tests by a lot of students. Those that use special pencils (and) answer sheets...are stressful; standardized tests are stressful. In (the lower grades) the students use the restroom during the test even though I take them before. To some kids, they get anxious not being able to sit through it. All of us feel 'tight' after the testing and try to make it an easier, less stressful activity.

Another second-grade teacher agreed. Her students, at a testing period, "cry, sigh, tap feet...(and) show relief when it's over." And one of the fifth-grade teachers was even more forceful in her description of negative test effects:

> The CTBS makes students act high for the rest of the day. Behavior is terrible afterwards. Even on local tests they will act up...They are louder, more uncon- trollable, (they) fight sometimes in the play ground (and find it) hard to sit still in a lot of situations if (the test) is too hard for them, like most tests are.

Another fifth-grade teacher agreed, though less vociferously, by describing her test-taking students as "drumming on the desk with pencils, fidgeting, and causing minor disturbances."

Issues stemming from student anxiety in the face of a need for testing were summarized by the school psychologist at Cityside. She noted that "some students get frustrated by some tests." Yet the same time, she recognized, that it may be difficult to "give a real good assessment without using a test. It's dificult; you need some kind of objective criteria."

Other commentary by Cityside teachers indicated that they were less concerned about testing's effects, on themselves and their students. These teachers believed that the more positive approach they took to testing made a difference. For example, one of the kindergarten teachers described the situation in these terms:

> Testing is a tool for me and not viewed as a burden. I just keep recycling. Tests that I give don't bother (the students) at all because I enjoy giving them and they're fun. I make (the students) absolutely aware that we're trying to find out something and that I need some information. I don't allow the students to get uptight.

This approach to test and testing is alluded to by several other teachers at Cityside. For example, a sixth-grade teacher mentioned that "test preparation is fundamental with our children."

That teacher attitude toward tests and testing varied within Cityside, and that this teacher attitude may have a bearing on the amount of stress felt by the students, was corroborated by the school's Title I Program Coordinator. According to this administrator, some teachers don't understand what a test is for or what the scores mean. Therefore:

> they'd complain and som wouldn't put forth the effort to make sure (they understand the test purpose). They'd give (the test) to the children and tell them to do the best they could.

The vice principal then went on to describe the ideal situation and practice which some of the teachers at Cityside try to follow. That is:

> ...to prepare (students) with the (testing) mechanics; not the test, but the mechanics

so that students understand how to take the test. This, the Coordinator said, can lead to improved student attitude and students' higher expectations for themselves.

At Hillview, all of the teachers referred in some manner to the cost that test-anxiety incurs for students. Taken jointly, these teacher comments suggested that while testing does not impose a uniformly high psychological stress for all students at Hillview. Nevertheless, comments reveal, <u>some</u> students do occasionally become over-anxious. For example, as explained by the kindergarten teacher at Hillview, "some kids feel pressured in the beginning (but) most kids are okay by May."

However, a first-grade teacher explained that:

> This is a highly competitive group of children. They know what group everyone's in and who's high and who's low--and we never mention it. And when a mastery test is given and we can't let some children go on to the next group, it's devastating to them.

Comments by other teachers at Hillview, especially in the upper grades, suggest that test anxiety does not apply to all students. Their remarks indicate that anxiety which does occur is usually manifested during curriculum or placement tests, which affect student standing in the classroom or placement in a subsequent grade or school. Less anxiety, in these teachers' view, appears during standardized tests which are not used for placement or promotion purposes at Hillview.

Pressure on Hillview students is also increased to some extent, staff members believed, because of parental influence. As a fifth-grade teacher put it:

> There's considerable parent pressure, particularly among Asian parents--a drive for students to get ahead. Parents will drop in and check how their child is doing. They will sign their children up for all different kinds of lessons. In many cases the children don't play with others.

Beyond the question of the anxiety instilled in students because of test or test-related pressures, the teachers at Cityside (but not at Hillview) made comments on other more positive effects of testing.

Student motivation; Cityside: Three or four teachers at Cityside cited testing as a reinforcer or motivator. According to a first-grade teacher:

> Testing is anxiety; that's a built in. That's part of life because you're being tested all the time. Actually that's probably good for (the students)...Once you overcome it and do it, next time you may be anxious but you know you can do it.

The sixth-grade teacher who had commented that test preparation is, or should be, fundamental at Cityside, agreed:

> I feel comfortable about tests. Kids need a certain amount of anxiety. There are no particular tests that cause my students anxiety.

This teacher then described her students' enjoyment and motivation from some kinds of tests:

> They get their (teacher-made spelling tests) back the same day. They love that. They always want to see how they did. They'll come to the aide or me and ask: 'Did you score the papers? Are they ready, yet?'

Obstacles to motivation; Cityside: Even Cityside teachers who would like to use tests as instructional motivators, However, found that there were obstacles to doing so. Describing the MAT, for instance, one of the fourth-grade teachers was disturbed that "students come out particularly low." Further, for formal tests in general, teachers may not agree with the accuracy of the results, because:

> Many times (the students) don't do well on paper-and-pencil tests. A lot is a guess. If they don't look, they make a mistake...Students may not be motivated. Most of the class has lots of family problems, and other things make it difficult for them. (This leads to) two extremes of (of test behavior); 'I can't do it' or 'I won't do it.' Then they give up.

The problem of students "giving up" was reiterated by the Title I Coordinator in terms that hark back to an earlier concern with test validity in general. That is:

> There are things in the CTBS that (some) children never come in contact with (and so) it's a waste of time. I think it's better if (the test) includes most of the things they come in contact with. And I think they are frustrated. They don't know the answers.

On the other hand, as indicated previously, it is possible in teachers' views for a student to get a false sense of accomplishment on the basis of scores

on tests like the District Continuum-Based Skills Survey. Because the ceiling on this test is so low, remarked one of the second-grade teachers (#13), the student "can have a good score and know nothing." The Title I Coordinator agreed: "(the Skills Survey) only has the minimum. Children can't be challenged if your expectations are the minimum."

The failure, or in some cases, inability, to use tests as instructional motivators was aptly decribed by the Bilingual Coordinator. According to this specialist, some students viewed the CTBS as a

> pass or fail situation, and therefore take that quite seriously. This is too bad. Student motivation is wasted because the test is used only for external (reporting) requirements.

Pressure growing from public reporting of scores; Hillview: The four teachers at Hillview commenting on this issue suggested that they are concerned that school administrators and the public believe that state- and district-mandated tests reflect teachers' work and therefore their competence. As a fourth-grade teacher at Hillview put it: "Handing in test results to the principal adds pressure." As explained by a fifth-grade colleague, "turning in test scores exerts a psychological pressure on the teacher because each spring the principal posts the standardized test scores by classroom," and "I think there's some pressure on teachers as a result of that." Further, according to this teacher, the principal had been stressing

that "he wants to know why" there has been a decline in primary-grade test scores, "and I think this creates some (teacher) anxiety."

How this kind of teacher anxiety in Hillview can grow was explained as follows by a first-grade teacher:

> I think that any time a test is given, a national type test, you don't lose sleep over it or anything, but you're concerned because it is your children being tested. There-fore it's what you have taught them and it is published and it is reflected back onto you if the students are below where they should be.

A fifth-grade colleague agreed:

> ...I would say there's a certain amount of pressure, not on the weekly or unit tests, but (on the) mandated tests at the end of the year...What our principal does is post a list of how the various classes have done. He makes it anonymous but we can figure it out. .....it would be very upsetting knowing that it's not always the teaching that produces that kind of score (a low growth score)...and sometimes you look at that kind of list and you know that other people are saying 'here's the good teacher and here's the bad teacher.' It's ludicous. I don't like that kind of comparison.

Loss of Instruction Time

While only one or two teachers at Cityside explicitly stated a concern with the intrusion of tests on instructional time, about half of the teachers at Hillview expressed this concern. As a first-grade teacher at Hillview put it, "testing cuts in on instructional time; for example students don't get reading instruction for two weeks." Her team-teaching colleague agreed that "tests add more work" and "cut instructional time."

Many teachers also indicated that some tests create behavior problems with students; hence (as described above) teachers routinely give over at least fifteen minutes of potential instructional time to allow students to wind down before resuming teaching-learning activities.

Summary

Teachers' commentary on psychological and other costs associated with testing generally reflected concerns with test utility or usefulness, the appropriateness of tests for students and/or the appropriateness of how their results are used, the effects of testing, and loss of instructional time caused by testing.

While these concerns were evident to some degree in both schools, the pattern of responses and emphasis varied. The Cityside data suggests that teachers were annoyed and somewhat frustrated with the imposition of tests that have limited utility and/or are of questionable worth and suitability in context. However, while they are a bit concerned about the anxiety that tests may cause students, tests are not viewed as a serious source of personal stress. Testing, in other words, may entail noteworthy opportunity costs in terms of time spent in useless or invalid pursuits, but significant psychological costs do not accrue.

In contrast, teachers at Hillview are more vocal about direct psychological costs of testing. All noted test-related anxiety in their students, and over half felt personally (albeit minimally) stressed and pressured by testing. These anxieties may result because test scores have both credibility and utility at Hillview--within an accountability context--for everyone in the setting. They carry personal consequences for both students and teachers.

PSYCHOLOGICAL COSTS: STUDENT ATTITUDES TOWARD TESTING

Relatively little is known about students' attitudes and feelings toward assessment in general. Even less is known regarding their feelings about different forms of assessment. In a 1979 study, Stetz and Beck asked students to respond about testing on a questionnaire consisting of semantic differential scales, e.g., helpful-harmful, unbiased-biased, calm-anxious, and supportive-antagonistic. At the K - 4 levels, a majority of students felt somewhat positively toward tests, although 56 percent indicated that they were nervous about taking them. At higher grade levels (5 - 12), only 26 percent of the students felt positively about tests, while 27 percent reported feeling negatively about them. In addition, 30 percent reported getting nervous before taking tests made by the teacher.

In a study by Sharp (1966) of 25 elementary and secondary teachers in Florida, there was an evenly mixed reaction to the question of whether emphasis on testing caused competitiveness in the classroom.

The question of whether test scores affect a student's self-concept has also been raised. Kirkland (1971) pointed out that the effect of receiving information about one's abilities will depend on a variety of factors, including the legitimacy of the source of the information, the perceived accuracy of the test, the degree to which the information confirms one's own estimate, and the extent to which it is threatening or rewarding. Test scores have potentially great impact where an individual's self-concept is at considerable variance with the record of performance on the test, where rationalizations of

poor peformance are unavailable, or where the test score is substantially higher than one's own estimate. Under such conditions, one can expect a shift to affect the individual's aspiration level, motivation to achieve, and personal decisions about the future. However, data from a national sample (Kirkland, 1971) indicated that test scores are of relatively minor importance in shaping one's self-estimate of ability in comparison with school grades, comments made by peers and parents, and a student's relationship with his/her teachers. But, Kirkland also reported that a majority of parents surveyed felt that their lives had been influenced by test results.

In light of these few and certainly non-definitive findings, student interviews were undertaken to explore the affective valence that different forms of achievement assessment have for students. Do they find testing a positive or negative experience? How worrisome do they find more and less formal means of assessment? How does the experience of assessment seem to influence their feelings about their own intelligence, and how others view them? How does the experience of assessment affect students' views about "what's important" in their academic career?

A three-part student interview schedule was developed to gauge students' responses to these and other questions about testing activities.

Interview Procedure

A systematic random sample of 60 students was selected from alphabetized class lists in the two case-study schools, Hillview and Cityside. The students were selected from the fourth, fifth, and

sixth grades at each school, totalling 20 students per grade level -- 10 each grade from the two schools. Included in the total sample were 37 males and 23 females. The overall ethnic composition of the group (using categories applied by the schools) was as follows: 26 Black; 13 White/Anglo; 6 Hispanic; 14 Asian; and 1 Pacific Islander.

The Interview Schedule

The interview was developed in a game-like format involving three tasks. (Please refer to Appendix C for a sample of the interview.) The first activity consisted of a sorting task called "Pick-Up-Sticks". The subject was asked to sort 10 common school activities, including six achievement-assessment activities, into 3 piles: "Activities I like": "Activities I dislike": and "Activities in the middle/no opinion". After this initial sort, the subject was asked to rank the activities in the "like" and "dislike" piles, putting the most liked (or most disliked) activity on top, followed by the next most liked (disliked), and so forth.

The second task involved a semantic differential exercise with 4 pairs of descriptors on a 7 point scale. Subjects were asked to place each of the ten school activities manipulated in Task #1 along the 7 point scale on each of the four semantic scales. (The scales themselves are described below.)

In the final task, students were asked to estimate which of 5 school assessment activities parents, teachers, they themselves, and their classmates thought that it was "most important to do well on."

There were several reasons for the structure of this instrument. First, the interview embedded various forms of assessment (standardized tests; chapter tests; and teacher-made quizzes; homework,

answering teachers' classroom questions, and story writing) amidst other forms of school activities. physical education games; assemblies; nutrition or snack time; talking with friends. The purpose of this was simply to see whether subjects did differentiate assessment from non-assessment activities, as well as to see whether students differentiated among different forms of assessment. Second, student attitudes toward the same testing and school activities were measured in three different ways. This not only provided a measure of the instruments' inherent construct validity, but also measured consistency of students' opinions across different elicitation contexts.

Administration Circumstances and Process

The instrument was administered individually to students in a quiet corner of the library or in an otherwise unoccupied resource room. In all cases, staff members and other students were either absent or well out of earshot during the interview.

After the interviewer introduced him/herself, he or she briefly explained that "we're talking to kids in lots of different schools about how they feel about different school activities." The interviewer emphasized that "there are no right or wrong answers" and that the talk was confidential, then proceeded to explain the first task. As the interviewer explained the task, s/he displayed the "game pieces." After asking any questions, the student was asked to do a sample item. The actual interview did not begin until the student demonstrated that s/he clearly understood what s/he was to do. However, students rarely had to repeat an example.

The game was already set up on one or two tables before each student arrived. For the first task, 3 large (7x4) index cards were

placed in a row. The cards were printed with the following: LIKE: IN THE MIDDLE/NO OPINION: DISLIKE. The student was then given the "sticks," tongue depressors, on which an activity was clearly marked in red. After the student had sorted and ranked these activities, s/he proceeded to the next task. Each task was preceeded by an explanation and a sample item.

For the second task, the game pieces were also displayed. These consisted of a number line marked from 1 to 7 and large index cards on either side of the number line. These cards were marked with the semantic differential descriptors. Using the same sticks s/he used for task 1, the student had to place or point each stick on the number line for each differential pair: fun/not fun, important/unimportant, smart/dumb, and calm/worried.

For the final task, the student was presented with a square divided into 20 cells. On the uppermost part of the figure five activities were listed (homework, teachers questions in class; standardized tests; chapter tests, and teacher made tests). On the vertical side of the figure the following were listed: my teacher; my folks; me; kids in my class. As the student answered the question, which activity would (your folks, kids in your class, etc.) like to see you do best on, the interviewer marked the appropriate cell.

This instrument was piloted on six successive occasions on a sample of 30 students at three elementary schools. The instrument was revised after each pilot occasion. The final pilot was performed with the instrument which was used in the study. The time for administration in the pilot and the study was from fifteen to twenty minutes per student.

Most students seemed to be quite comfortable with this instrument and understood the directions easily. As might be expected, older students finished the instrument a bit more quickly and often preferred to point or answer verbally rather than to manipulate sticks. All items were read and repeated to students to avoid interference of reading comprehension or other skills with the task.

III. The Findings

The subsequent sections report the findings from 3 perspectives.

First, we discuss student ratings on the importance of testing activities on tasks 2 and 3 (semantic differential and important-to-do-well-on). These findings indicate the importance of different types of testing; testing as compared to non-testing activities; and the realtionship between assessment activities and significant others in the eyes of the student.

The second perspective provides students' global affective responses to different types of assessment activities based on the like/dislike task.

The third section provides a more differentiated look at student feelings about assessment compared with other school activities.

Students' Views of the Relative Importance of Different Types of Assessment

A first issue was whether students considered various types of assessment of different importance. Thus, as we mentioned previously, six commonly used forms of students assessment were included in all three tasks on the instrument. These were chapter tests, standardized tests, teacher made quizzes, homework, writing a story,

writing a story, and answering teacher's questions in class. Notice that the first three assessment types are more formal, less frequent, and more clearly "marked" as instances of assessment. The other usually occur more frequently as part of the regular school routine and/or as more or less formal ways of evaluating students' achievement.

In addition to the six assessment modes, four other school activities were included in two of the tasks on the measure. These included recess, talking to friends, p.e. games, and assemblies.

Table 45 below illustrates that students regard assessment activities as more important than non-assessment activities. Clearly, standardized tests and chapter tests were rated as the most important activities. Assemblies (a non-assessment activity) were viewed as slightly more important than writing a story, which many teachers use to assess language arts skills. (Students may associate assemblies with instruction; assemblies in these schools are often used to convey information about school rules and regulations and to show educational films.)

Student ratings on the "important to do well on" task generally supported these findings (see Table 46 below).

Table 45

Overall Sample: Ordered Mean Ratings for 10 School Activities Important/ Unimportant (n = 60)

Standard-ized Test	Chapter Test	Home-work	Answering Teacher's Questions	Teacher Quiz	Assemblies	Writing A Story	P.E. Games	Recess/ Nutrition	Talking With Friends
6.63	6.15	6.08	5.80	5.68	5.43	5.33	5.28	4.71	4.41

Table 46

Overall Sample: Frequency of Ratings on "Most Important to Do Well On" Task (n = 60)

	Home-work	Answer Teacher's Questions	Standard-ized Test	Chapter Test	Teacher Made Quiz
My Teacher	20%	5%	52%	17%	5%
My Folks	40%	7%	33%	10%	8%
Me	17%	12%	43%	20%	7%
Kids in My Class	13%	18%	22%	22%	22%

Over half the student sample (52%) responded that teachers feel it is most important to do well on standardized tests. About 43% of the students also named the standardized test as the assessment type that they themselves believed it was most important to do well on. The sample was closely dividerd with regard to parental views: 40% said parents would rate homework as the most important and 33% indicated that standardized tests would be the parents' choice.

Although students in both schools gave standardized tests a similarly high rating across all Significant Others, there were some differences with respect to other activities. Cityside students indicated that they and their teachers would consider homework to be the next most important activity. Hillview students, on the other hand, rated chapter tests as the next most important. This pattern is also repeated in Table 48 below, which shows between-school differences in their ranking of assessment activities. Note also that Hillview students rated writing a story as much less important than did students at Cityside.

Table 43

Frequency of Rating for "Most Important to Do Well On" Task by School
[Cityside, n = 30; Hillview, n = 30]

	Homework		Answering Teacher's Questions		Standardized Test		Chapter Test		Teacher Made Quiz	
	City-side	Hill-view	City-side	Hill-view	City-side	Hill-view	City-side	Hill-view	City-side	Hill-view
My Teacher	8	4	2	1	16	15	2	8	1	2
My Folks	12	12	2	2	10	10	2	4	3	2
Me	7	3	5	2	12	14	5	7	--	4
Kids in My Class	3	5	7	4	7	6	6	7	5	8

Table 48

Mean Rating for Assessment Activities by School: Important/Unimportant
[Cityside, n = 30; Hillview, n = 30]

	Standard-ized Test	Home-work	Chapter Test	Answering Teacher's Questions	Teacher Made Quiz	Writing A Story
Cityside	6.73	6.43	6.23	6.03	5.86	5.86

	Standard-ized Test	Chapter Test	Home-work	Answering Teacher's Questions	Teacher Made Quiz	Writing A Story
Hillview	6.53	6.06	5.73	5.56	5.50	4.80

Table 49 displays students' mean ratings on the "importance" semantic scale by grade level. Across all three, students rated standardized tests as the most important activity. Chapter tests and Homework continue to stand out as among the important forms of assessment, but notice that which is given priority alternates across grade level.

Notice too that mean reatings for all six assessment forms tend to decrease across the upper elementary grades. The small sample size (n = 20 per grade level) and degree of these differences suggest circumspect treatment. Perhaps, however, the differences reflect that students find the assessment experience - whatever its form - more routine and less awe-inspiring as they continue through school.

Table 49

Mean Rating for Assessment Activities by Grade: Important/Unimportant
[Grade 4, n = 20; Grade 5, n = 20; Grade 6, n = 20]

	Home-work	Writing A Story	Standard-ized Test	Answering Teacher's Questions	Chapter Test	Teacher Made Quiz
Grade 4	6.30	5.60	6.65	6.05	6.50	6.15
Grade 5	6.20	5.30	6.80	5.70	6.15	5.75
Grade 6	5.75	5.10	6.45	5.65	5.80	5.15

In summary, the sixty students interviewed rated all six assessment modes on the "important" side of the semantic scale. Nevertheless, on the whole, they saw two more formal and (usually)

more comprehensive modes - standardized tests and chapter tests - as more important than the others. Homework (which many respondents believed their parents emphasized) was also given a comparatively high importance rating across two interview tasks. Routine oral evaluation (answering classroom questions) and quizzes followed in close succession. Thus, students' mean ratings of importance seem in a general way to reflect the following principle: measures that occur less frequently and "cover" more content tend to be more important. And in practice, measures of that kind do very often weigh more heavily in evaluating student performance.

B. Students' General Demeanor Toward Different Forms of Assessment

The foregoing discussion describes part of students' conceptualizations of classroom assessment activities. It suggests that at least by the upper elementary grades, pupils can and do differentiate among the relative importance of different forms of assessment. Broadly speaking, their views seem consonant with actual practice. Each instance of a standardized test or a chapter test usually has the potential of making more difference in students' educational careers than each instance of a quiz, homework, or oral classroom performance.

A second issue which seemed worth exploring was students' general affective demeanor toward assessment, and whether their general feelings vary with different types of assessment techniques. The sorting task described previously attempted to examine this aspect of students' attitude.

To review, students were asked to sort the some ten activities just discussed including the six forms of assessment into three piles:

"things I like," "things I dislike," and "things in the middle." They were then asked to rank order the activities placed in the "like" and dislike" piles.

As might be expected, students consistently preferred the non academic (53%-93%) to the assessment activities. (See Table 50.) The next most liked activities, overall, were the more routine, less marked forms of assessment (32-57%). Direct testing activities were less often mentioned as liked (17-38%). Conversely, the most disliked activities were usually the direct forms of testing (20-43%), followed by indirect assessment activities (17-30%) and social school activities (3-8%). It should be noted that a significant percentage of the sixty students (23-42%) took a "neutral" position on the appeal of assessment, placing various modes "in the middle."

Table 50

Percentage of Students Who Labeled Each School Activity as "Like", "In the Middle", or "Dislike": Total for Both Schools

	LIKE	IN THE MIDDLE	DISLIKE	TOTAL
Standardized Tests	32%	27%	41%	100
Chapter Tests	17%	40%	43%	100
Teacher Made Quiz	38%	42%	20%	100
Homework	32%	38%	30%	100
Writing a Story	57%	23%	20%	100
Answering Questions	45%	38%	17%	100
Assemblies	53%	38%	9%	100
P.E.	87%	5%	8%	100
Recess	82%	15%	3%	100
Talking with Friends	93%	2%	5%	100

Three observations are worth making here. The types of assessment that students on the whole like less often and dislike more often are those that they collectively rated as more important: those that tend to be less frequently administered and more comprehensive in content (standardized and chapter tests), along with homework (which makes a regular claim on children's out-of-school time). Second, a majority of the students interviewed reported viewing even these performance modes positively or neutrally. And only small proportions of students reported disliking quizzes and answering teacher's questions, while more than half said they enjoyed writing a story. Nevertheless (third), the minority that expressed dislike for the less frequent, more formal and comprehensive forms of testing was a substantial one.

In Table 51, certain differences in student's attitudes are evident between schools. The most notable of these lies in students' preferences toward standardized tests: 53% of the students at Cityside said they liked standardized tests as opposed to only 10% of the students at Hillview. At the same time, 50% of the students at Hillview said they disliked these tests, compared to 30% at Cityside. The same pattern holds for chapter tests. And overall, at Hillview the frequency of like responses is lower for each academic assessment activity; Hillview students tend to be more affectively neutral on most.

Finally, it is worth underscoring that students at both schools, on the whole <u>did</u> offer differentiated responses on the sorting task. This is especially evident when their reactions to the academic school activities are compared to their reactions toward the non-academic ones.

TABLE 51

Percentage of Students Who Labeled Each School Activity as "Liked", "In the Middle", or "Disliked" Total by Schools

	KING			HILLVIEW		
	LIKE	MIDDLE	DISLIKE	LIKE	MIDDLE	DISLIKE
Standardized Tests	53%	17%	30%	10%	37%	53%
Chapter Tests	30	33	37	3	47	50
Teacher Made Quizzes	50	30	20	27	53	20
Homework	50	20	30	13	57	30
Writing a Story	60	7	33	53	40	7
Answering Teacher's Questions	60	23	17	30	53	17
Assemblies	43	44	13	64	33	3
P.E.	90	7	3	86	3	13
Recess	83	10	7	80	20	-
Talking with Friends	90	3	7	97	-	3

A Finer-Grained View of Students' Feelings About Testing

The results of the sort-and-rank task, just discussed, provide a look at students' global feelings toward different forms of assessment. In general (and especially at Hillview) the more formal and comprehensive tests - standardized and chapter - were viewed most negatively. But only about two-fifths of the interviewees found these unappealing, and a majority of responses to each assessment mode were positive and neutral.

Now, we turn to a more differentiated view of the positive and negative valence of assessment for students. In the semantic differ

ential task previously described, students were asked to place each of the six assessment and four non-academic activities on the following scales: (1) fun/not fun; (2) calm/worried; and (3) smart/dumb.*

1. Students' Experience of Different Assessment Forms as Fun or Not Fun

The fun/not fun scale probably taps an affective dimension similar to the "like to the middle of dislike" sorting task.** It goes beyond that task, however, in revealing the magnitude of individual students' general feelings about the different assessment modes.

As Table 52 shows, non-academic activities received higher mean ratings than the assessment activities. Once again, standardized tests, homework and chapeter tests were the most negatively rated.

Table 52

Overall Sample: Mean Ratings for 10 School Activities Fun/Not Fun (n = 60)

Standard-ized Test	Home-work	Chapter Test	Answering Teacher's Questions	Teacher Made Quiz	Assemblies	Writing A Story	P.E. Games	Talking With Friends	Recess/ Nutrition
3.50	4.06	4.08	4.88	4.96	5.00	5.16	6.30	6.31	6.43

* The result of students' responses on a fourth scale, important/unimportant, have already been discussed.

** A cross tabulation shows that, overall, individual students' responses on the sorting task were consonant with their ratings for the same items on the fun/not fun scale for 79% of the interviewees. A consonant response is defined broadly here as (1) a "like" placement on the sorting task with a rating of 7,6, or 5 on the seven-point fun/not fun scale; or (2) an "in the middle" placement woth a 5,4, or 3 rating; or (3) a "dislike" placement with a 1, 2, or 3 rating. This definition slightly braodens the "middle" range of semantic differential scale, which is of course constituted only by the rating "4".

However, Table 53 below, which describes the frequency of ratings for the six assessment items, shows that the sample was almost evenly divided on their ratings for some of the testing items.

Table 53

Overall Sample: Frequency of Ratings for 6 Assessment Activities Fun/Not Fun (n = 60)

	Fun						Not Fun
	7	6	5	4	3	2	1
Homework	20%	7%	20%	15%	10%	8%	20%
Writing a Story	37%	17%	17%	8%	7%	8%	7%
Standardized Test	15%	10%	8%	15%	15%	7%	30%
Answering Teacher's Questions	22%	18%	13%	32%	7%	5%	3%
Chapter Test	15%	13%	15%	17%	17%	8%	15%
Teacher-Made Quiz	30%	15%	15%	20%	7%	8%	5%

Only one activity, standardized tests, was negatively ranked by 50% or more of the sample. Although chapter tests and homework were negatively rated by 38 to 40% of the sample, they received positive ratings by 43 to 47% of the sample. Note too, that these items received distinctly higher percentages of ratings of "1", at the extreme negative end of the scale. Other assessment activities received more positive than low negative ratings. Writing a story was rated fun (5-7) by 71%; teacher-made quizzes by 60%; and answering teacher's questions in class by 53%.

The between school comparison of ratings seen below in Table 54 confirms patterns already described. That is, standardized tests, homework, and chapter tests are the most negatively rated activities by both schools. A significant means difference was found only for the teacher-made quiz, where Hillview students assigned a more negative rating ($p \leq .01$).

Table 54

Mean Ratings for 6 Assessment Activities by School
Fun/Not Fun

	Standard-ized Test	Chapter Test	Home-work	Writing A Story	Teacher Made Quiz**	Answering Teacher's Questions
Cityside	4.06	4.33	4.53	5.53	5.66	5.23

	Standard-ized Test	Home-work	Chapter Test	Teacher Made Quiz	Answering Teacher's Questions	Writing A Story
Hillview	3.03	3.60	3.83	4.26	4.53	4.80

Similar findings were found when grade level comparisons of ratings were done. As Table 55 below indicates, homework and standardized tests usually receive negative (less than 4) ratings whereas writing a story, answering teacher's questions and doing teacher-made quizzes receive positive (5 or more) or neutral (4) ratings.

Table 55

Mean Rating of 6 Assessment Activities at Three Grade Levels: Fun/Not Fun
[Grade 4, n = 20; Grade 5, n = 20; Grade 6, n = 20]

	Home-work	Writing A Story	Standard-ized Test	Answering Teacher's Questions	Chapter Test	Teacher Made Quiz
Grade 4	4.85	5.40	3.20	5.10	4.50	5.35
Grade 5	3.85	4.95	4.20	4.85	3.75	4.70
Grade 6	3.50	5.15	3.25	4.70	4.00	4.85

In summary, a majority of the students interviewed found three less-formal, more-routine forms of assessment to be fun. And the sample's mean responses confirm that for most pupils standardized tests, chapter tests, and homework are the least appealing forms of assessment. Finally, it is notable that roughly a quarter to a third of the students interviewed experience these activities as more-or-less aversive: about this proportion rates each with either a "1" or "2" at the negative end of the fun/not fun scale.

2. Students' Views of Different Forms of Assessment as Worrisome

To what extent do students seem to worry when confronted with different types of assessment?

The mean ratings for the overall sample (Table 56) shows that students feel calm in all non assessment items and in one assessment item, writing a story. Their ratings of other assessment items were neutral.

Table 56

Overall Sample -- Mean Rating for 10 School Activities
Calm/Worried (n = 60)

Standard-ized Test	Home-work	Answering Teacher's Questions	Chapter Test	Teacher-Made Quiz	Assemblies	Writing A Story	P.E. Games	Recess/ Nutrition	Talking With Friends
4.08	4.33	4.63	4.46	4.71	5.00	5.33	5.85	5.95	6.10

However, when we look at the frequency of ratings for the six assessment activities in Table 57 below, we find that a small though significant proportion of students, 26 to 38%, worry about some forms of assessment: standardized tests (38%); homework (34%); chapter tests (27%); and answering teacher's questions (26%). The greater proportion of students feel calm across all activities, particularly in writing a story (68%), taking a teacher-made quiz (59%), doing a chapter test (51%), and answering teacher's questions (50%).

Table 57

Overall Sample: Frequency of Ratings for 6 Assessment Activities Calm/Worried (n = 60)

	Calm						Worried
	7	6	5	4	3	2	1
Homework	17%	10%	17%	23%	20%	7%	7%
Writing a Story	33%	23%	12%	17%	7%	7%	2%
Standardized Test	15%	17%	7%	23%	13%	12%	13%
Answering Teacher's Questions	20%	12%	18%	23%	18%	5%	3%
Chapter Test	22%	17%	12%	23%	7%	3%	17%
Teacher-Made Quiz	17%	22%	20%	18%	13%	2%	8%

Between school ratings (Table 58) show only that students in both rated themselves calm in writing a story. The only school-to-school difference was that Hillview students gave homework a negative (worry) rating unlike Cityside. All other ratings were neutral.

Table 58

Mean Ratings for 6 Assessment Activities by School: Calm/Worried
[School 1, n = 30; School 2, n = 30]

	Standard-ized Test	Chapter Test	Teacher Made Quiz	Home-work	Answering Teacher's Questions	Writing A Story
Cityside	4.13	4.43	4.60	4.76	4.96	5.56

	Home-work	Standard-.ized Test	Answering Teacher's Questions	Chapter Test	Teacher Made Quiz	Writing A Story
Hillview	3.90	4.03	4.30	4.50	4.83	5.10

A display of mean responses on the calm/worried scale shows no general trends. Viewed in juxtaposition with Table 50, however, one minor point emerges. While students mean ratings of the <u>importance</u> of all assessment forms declines across grade levels, there is no accompanying decline in how much <u>worry</u> students associate with them.

Table 59

Mean Rating of 6 Assessment Activities at Three Grade Levels: Calm/Worried
[Grade 4, n = 19; Grade 5, n = 20; Grade 6, n = 20]

	Home-work	Writing A Story	Standard-ized Test	Answering Teacher's Questions	Chapter Test	Teacher Made Quiz
Grade 4	4.35	5.35	3.85	4.45	4.70	4.65
Grade 5	4.55	5.40	4.90	5.00	4.30	4.60
Grade 6	4.10	5.25	3.50	4.45	4.40	4.90

3. Students' Association of Forms of Assessment with Their Intellectual Self-Esteem

Assessment activities provide occasions for students to do well or poorly, to succeed or fail. Presumably, then, they can influence students' perceptions of their own intellectual competence. What kind of influence assessment has probably depends upon how well students perform when assessed. Nevertheless, it seemed worthwhile to explore the extent to which generic forms of assessment were associated for students with feelings of intellectual capability or incapability. The smart/dumb semantic scale was intended to examine this issue in a general way.

Overall, students did not differentiate the six assessment activities along the smart/dumb semantic scale. As Table 60 illustrates, the testing activities received ratings which ranged from a low of 5.36 to a high of 5.65 for the total sample (n = 60). These differences are significant neither intuitively nor statistically.

Table 60

Overall Sample: Ranked Mean Ratings for 6 School Assessment Activities Smart/Dumb (n = 60)

Standard-ized Test	Writing A Story	Teacher Made Quiz	Answering Teacher's Questions	Chapter Test	Home-work
5.36	5.55	5.55	5.60	5.65	5.70

The overall frequency of ratings for assessment the items (Table 61) shows that 68 to 83 percent of the responses was within the from 7 to 5 range (smart") for all items; 12 to 23 percent were in the exact middle of the scale; and only 2 to 8 percent on the negative ("dumb") side of the scale. (Also see mean ratings for each schools' students in Table 62.)

These findings may reflect a reluctance on students' parts to admit feeling "dumb", especially to a stranger. It may be, too, that the structure of this question was confusing: students may not have been able to associate a general view of themselves as feeling "smart" or "dumb" with a generic assessment activity. However, pilot interviews employing this same item "worked" to elicit a substantially wider range of responses. It may simply be, then, that - whatever their individual performance - students at Hillview and Cityside rarely felt very "dumb" in the mere presence of assessment activities.

Ethnographic work in the two schools (conducted in conjunction with this and earlier projects) suggests that teachers believe strongly that their students are capable. They appear to routinelycommunicate this to the children. Hillview is often spoken of in Littleton District as the school with the highest achievers. Cityside was recently cited as outstanding among the Metro District schools with compensatory education programs. Word of their schools' relative standings probably makes its way to students. And within each setting, most students progress through their subjects with rates of achievement that permit them to feel competent. Few are likely to receive consistent evidence that they are incapable academically. Their responses on the "smart/dumb" scale may very well reflect this demonstrable fact.

Table 61

Overall Sample: Frequency of Ratings for 6 Assessment Activities Smart/Dumb (n = 60)

	Smart						Dumb
	7	6	5	4	3	2	1
Homework	40%	22%	13%	20%	3%	2%	—
Writing a Story	38%	18%	12%	23%	8%	—	—
Standardized Test	37%	15%	20%	17%	3%	5%	3%
Answering Teacher's Questions	37%	22%	17%	18%	3%	3%	—
Chapter Test	33%	20%	28%	12%	2%	3%	—
Teacher-Made Quiz	25%	32%	22%	18%	2%	2%	—

Table 62

Mean Ratings for 6 Assessment Activities by School: Smart/Dumb [Hillview, n = 30; Cityside, n = 30]

	Teacher-Made Quiz	Writing A Story	Standardized Test	Chapter Test	Answering Teacher's Questions	Homework
Cityside	5.76	5.93	6.00	6.00	5.93	6.36

	Standardized Test	Homework	Writing A Story	Answering Teacher's Questions	Chapter Test	Teacher-Made Quiz
Hillview	4.73	5.03	5.16	5.26	5.30	5.33

D. Summary

The data show that students distinguish assessment from non assessment activities across all tasks, and within assessment items on some. Students rated standardized tests as the most important and worrisome activity as well as among the least liked and least fun. Chapter tests and homework competed for second place as the most important, least liked and least fun activity. Their second place rating varied according to whether responses were examined for the total sample, by school, or across grade levels. Teacher made quizzes and answering teacher's questions in class also vied for third place in importance. However, students usually rated them likeable and fun activities. The most popular assessment activity was writing a story. It was given the highest fun and like ratings of the six assessment activities. It was also rated to be the least important one.

The general between-school pattern across the instrument is that Cityside students gave slightly to moderately higher (positive) ratings than Hillview students did on the "like/dislike" tasks and "fun/not fun" scale.

Across-grade-level variations showed a slight trend: attitudes toward standardized testing, chapter tests, and homework seemed to be more negative in higher grade levels. These activities were experienced as less liked, less fun, and more worrisome by the sixth graders than by the fourth graders. It is interesting to note that these as well as other assessment activities, were viewed as less important from the fourth to the sixth grade.

Student ratings on the dimensions of affect (fun/not fun, calm/worried, smart/dumb) support teachers' comments on the psychological costs of testing. Teachers indicated that although the majority of their students did not find most assessment activities to be a particularly worrisome or negative experience, a minority of students did manifest anxiety by complaining or, in a few instances, crying. Most students indicated that they felt calm and smart during all testing activities even though they did not rate them as fun activities. This includes those activities rated as very important. However, about one third or more of the students (38 to 40%) expressed feelings of anxiety or distaste for standardized tests and chapter tests.

Because of the small sample size (n = 60) and the paucity of research in this topic, these findings suggests potential avenues for research as much as they provide information. For example, Cityside students had generally more positive attitudes toward testing than did Hillview students. Recall that Cityside is an inner city moderate to low income school. This finding contradicts the stereotypical notion that inner city students are less self-confident and receptive toward testing than their middle class fellow students in the suburbs, such as Hillview. However, further studies with larger student samples would be needed in order to validate this finding.

Students in both schools seemed to find teacher-oriented activities (i.e. quizzes, class questions, story writing) much more positive than the more formal and less frequent standardized tests and chapter tests. It would be interesting and useful (for instructional purposes) to ascertain whether the frequency and source of a tests as

well as its potential effect on a student's career, influence their motivation and attitude toward assessment.

Ratings toward writing a story are also worth exploring. This assessment technique was thought to be the least important though the most fun and best liked activity. Did students consider this to be an assessment activity or an instructional technique? Had they been asked for their ratings on writing an essay in science or history, would their ratings have changed?

These findings and the issues they raise make evident the need for further research and perhaps a rethinking of current notions about student attitudes toward testing.

Teacher and Student Commentaries on the
Psychological Costs of Assessment: A Summary

The teacher and student interviews which examined the psychological effects of assessment support one another on several points.

Overall, teacher and student interviews suggest that tests are not a source of serious stress for most students. However, for a minority of students, testing can be stressful.

The findings also indicate that tests which occur less frequently and which may seem to have broader impact on school careers (i.e. standardized tests and competency tests) are a somewhat greater source of stress than the more routine and perhaps less momentous tests such as teacher-made quizzes. Both teachers' comments and students' responses point to standardized tests as slightly to moderately stressful for students.

However, teachers and students seemed to disagree on one point. Some teachers claimed that unit tests (i.e. chapter tests, mastery tests) were not a source of anxiety. Most teachers did not mention this type of test in relation to their frustrations or aggravations with testing. On the other hand, students regarded chapter tests as the next most important and stressful type of assessment after standardized tests. Students in Hillview School also said homework could be worrisome, yet teachers did not comment at all on homework.

Students indicated that they viewed standardized tests, chapter tests, and homework as the most important assessment activities (in this rank order). They also suggested that their teachers would agree that these activities are the most important for students to do well on, perhaps a misperception, given teacher comments about the utility and appropriateness of the standardized tests that they gave.

Students on the whole reacted positively (on the like/dislike and fun/not fun scales) to teacher-made quizzes, answering teacher's questions in class, and writing a story, all instructionally related forms of assessment. Students also indicated that these were the least important forms of assessment, perhaps because they affect students' schooling in a cumulative rather than in an immediate or abrupt manner. Whereas students are aware that standardized tests and chapter tests examine a large body of knowledge, and will have an effect in their placement within the classroom, school, or future schools (i.e. junior high placement), more routine tests may not seem to have an effect on these aspects of a students' career. Teacher comments from Hillside support this. District tests, such as the District-mandated math operations test or the fourth-grade proficiency tests seem to cause more anxiety than the standardized tests. Results for the operations and proficiency test are posted. Awards are handed out for high achievement in the math operations test. Students who have not achieved high scores on this test exhibit keen disappointment, according to teachers. There are explicit and public consequences to performance on some tests, and these consequences may be a significant determiner of the psychological costs associated with testing.

To summarize, students and teachers did not indicate that assessment causes great anxiety. However, both agree that standardized tests and competency tests cause more stress than other forms of assessment. Assessment which is more narrowly related to instruction or the daily routine, seems to cause little stress. In fact, both teachers and students provided positive comments about these forms of assessment.

From these findings, we can speculate that at the elementary level stress arises from the prospect of being judged by peers and superiors (as in the case of Hillview), or from the frustration of coping with instructionally unrelated tests (as in Cityside's case). The impression that the less frequent tests (standardized and proficiency tests) have greater impact than the routine tests (such as spelling tests) may also be a source of anxiety.

APPENDIX A

TEST USE PHASE II

Teacher Questionnaire

Introduction

Before we begin, let me tell you something about who I am and the purpose of our interview today.

I'm _____ from the School of Education at UCLA, and specifically do my work at a research laboratory called the Center for the Study of Evaluation (CSE).

We're here in (name of district/school) as part of a three-year, national study that we started in 1979, so now we're in the final year. Let me tell you a little about that project. Basically, the first part of the study has been finding out about the many different ways that teachers and others go about assessing students' performance and progress. This can be a very complex process, and we have always felt that teachers have many good and useful ways of doing it. But back in '79 it was becoming clear that although a lot was said about how teachers make assessment decisions about students, very little of the information used to make these statements actually came from the teachers themselves.

To get as full a picture as possible of how teachers make assessment decisions, we decided to focus our study on all the ways that teachers have for making decisions about their students: from large-scale commercially published tests like CTBS, the IOWA, the SAT, and so forth, to other kinds of tests like those that come with textbooks, to ones that the district or that teachers make up themselves, and to other important kinds of information like teachers' classroom observations and use of professional judgment. In the past two years we've started to get a clear picture of how teachers use these various assessment techniques in their classrooms.

In this second part of the study, our job is just as important as the first part. Now, we're trying to get an accurate picture about how much time it all takes, and again we want to get that information directly from teachers.

Now, I'll get back to this later, but let me mention that just as we are interested in the total range of assessment techniques you use in your classrooms, or that others use with your students, we're also interested in the different ways that assessment takes up time, and therefore has a cost. First of all, let's consider the time that you, your students, and others put into testing and test-related activities. Every time you do something directly on or related to testing, there is some kind of monetary cost; every time you do something on testing, you have to give up the opportunity to do something else. You might have thought of some other ways to use the time had you not been testing. Finally, some testing activities may have a psychological impact.

Anyway, that's the project in a nutshell. Is there anything you'd like me to clarify before we go on?

Let me emphasize that your participation is voluntary and that each person included in the study will remain anonymous.

Any questions about any of that?

Tape Recording

Now, since I don't want to miss any of what you say, or inadvertantly change your words, I'd like your permission to tape record our talk. No identifying details will appear on the tape label, and only our project staff will be allowed to hear the tape if they need to transcribe it. If at any point you want to turn the recorder off, you just need to press this button. (DEMONSTRATE)

So is it okay if we tape record?

Let's begin with some background information.

I. Before we start exploring the testing issue, I would like some background information to get an idea of the context in which the testing situations occur.

1. First, I'd like to know about what grade(s) you teach.

2. Besides teaching, do you have any other responsibilities here?

3. How long have you been teaching at (name of school)? How long altogether?

4. Are the students in (specify the class grade) any particular tracks or ability groups? (If teacher needs clarification, provide terms such as: low, middle, high, regular, gifted, cross grade, etc.)

5. Is there an aide who works with the students in this class?

6. Is there a specialist who works with students in this class?

7. Do you do your teaching in any kind of a team arrangement?

II. Okay. Thank you. Let me briefly describe how we will proceed. Let's begin with those tests that are given infrequently, perhaps only at the beginning and/or end of the school year. Then, we'll talk about tests that you give routinely through the year, say, once a month or every couple of months, or once a week. Finally, we'll talk about those you use on a daily basis. We'll talk about each, from the least to most frequently given tests, in terms of the preplanning sheet you received.

CORE QUESTIONS

1. What kinds of tests are given on a _____ (supply time frame) basis in reading, language arts, math, science, social studies, and general achievement? (Get subject, test name, and test type.)

   PROBE FOR YEARLY: Have we covered all the tests that occur on a yearly basis? For example, competency tests, placement tests, or required pre and post tests?

   PROBE FOR MONTHLY: What about midterms, end of unit/book tests?

   PROBE FOR WEEKLY: What about book reports, compositions, or spelling and math tests/quizzes?

   PROBE FOR DAILY: What about questions at the end of a story or chapter? Do you ask questions reviewing previous work?

2. Does anyone make you give this test? If so, who?

3. Approximately when is this test given during the year? That is, approximate months or points during the year?

4. How many times is the _____(name test) given to the typical student during the year?

5. How much time and whose time is used in activities before, during and after administration? For example, there could be the time taken to construct the test or quiz, going to meetings to discuss how to administer the test, or preparing materials for the test, all before you actually administer it. During the test there is its actual administration, or having an aide act as proctor. After the test you might need to score it, review answers with students, and so forth.

Probe:	Before	During	After
	Test construction	Setting up	Scoring
	Informing students	Administering test	Grading
	Preparing materials	Proctoring	Interpreting
	Inservice activities		Reviewing

NOTE TO INTERVIEWER: Please refer to the corresponding worksheet as you ask the core questions and go through the appropriate routine.

6. Do you feel that the amount of testing you do overall is representative of the amount of testing that most teachers do in this school?

   Probe: Do most teachers spend as much time in testing math, reading, etc?

7. Do you do more testing in one particular area than most teachers in your school?

   Probe: For example, do you do more testing in reading (or other subject) than other teachers?

8. Of the tests that you give, are there any that you would eliminate? Which ones?

9. Other than the tests you have just told me about, do you have other ways of getting information about your students (Information from cum file, past teacher records, book reports.) How much time is spent doing this?

10. Are there certain kinds of tests that provide you (the teacher) with particular anxieties or stresses and concerns that make your work more difficult?

    Probe: One of the things that we are trying to do is to identify the "psychological costs of testing". What would you say are the psychological cost of testing? (For example are there changes in lessons or styles of teaching or anxieties over teacher evaluations.)

11. Are there particular tests that cause stress or anxiety to your students?

    Probe: How does that manifest itself? Are there other psychological costs of testing for students? (For example, misplacement, dropout, parental conflict.)

12. How and to whom are your concerns voiced?

13. Any other problems, difficulties and concerns for you or anyone else connected with the business of testing?

DATA RECORDING SHEET - (Teachers)

INTERVIEWER _____

I. Background

1. Grades _____
2. Other responsibilities _____
3. Time at school _____

   Total _____
4. Ability groups _____

   _____
5. Aides? _____
6. Specialists _____
7. Team teaching _____

GO TO TEST SHEETS

TEACHER RECORDING CHART

1. Subject _____
   Test _____
   Type _____

2. Who says _____
3. When given _____
4. X per year _____

	WHAT	WHO (circle as apply)	TIME
5.	_____	T S PV A C	_____
B	_____	T S PV A C	_____
E	_____	T S PV A C	_____
F	_____	T S PV A C	_____
O	_____	T S PV A C	_____
R	_____	T S PV A C	_____
E	_____	T S PV A C	_____
	_____	T S PV A C	_____
	_____	T S PV A C	_____
D	_____	T S PV A C	_____
U	_____	T S PV A C	_____
R	_____	T S PV A C	_____
I	_____	T S PV A C	_____
N	_____	T S PV A C	_____
G	_____	T S PV A C	_____
	_____	T S PV A C	_____
	_____	T S PV A C	_____
	_____	T S PV A C	_____
A	_____	T S PV A C	_____
F	_____	T S PV A C	_____
T	_____	T S PV A C	_____
E	_____	T S PV A C	_____
R	_____	T S PV A C	_____
	_____	T S PV A C	_____

T = Teacher S = Student PV = Parent Volunteer A = Aide C = Clerical

6. Teacher's testing is representative: YES _____ NO _____

7. More testing in one area: YES _____ NO _____

If yes, subject: _____

8. Tests to eliminate: YES _____ NO _____

If yes, what tests and why? _____

_____

_____

9. Information other than tests: YES _____ NO _____

If yes, what and why? _____

_____

_____

_____

10. Anxieties/Teacher

_____

_____

_____

11a. Anxieties/student

_____

_____

_____

11b. Manifestations

_____

_____

_____

12. Concerns voiced to:

_____

_____

_____

13. Other comments:

_____

_____

_____

_____

APPENDIX B

TEST USE PHASE II

Administrative Questionnaire

Introduction

Before we begin, let me tell you something about who I am and the purpose of our interview today.

I'm _____ from the School of Education at UCLA, and specifically do my work at a research laboratory called the Center for the Study of Evaluation (CSE).

We're here in (name of district/school) as part of a three-year, national study that we started in 1979, so now we're in the final year. Let me tell you a little about that project. Basically, the first part of the study has been finding out about the many different ways that teachers and others go about assessing students' performance and progress. This can be a very complex process, and we have always felt that teachers have many good and useful ways of doing it. But back in 1979 it was becoming clear that although a lot was said about how teachers make assessment decisions about students, very little of the information used to make these statements actually came from the teachers and administrators.

To get as full a picture as possible of how administrators and teachers make assessment decisions, we decided to focus our study on all the ways that teachers have for making decisions about their students: from large-scale commercially published tests like CTBS, the IOWA, the SAT, and so forth, to other kinds of tests like those that come with textbooks, to ones that the district or that teachers make up themselves, and to other important kinds of information like teachers' classroom observations and use of professional judgment. In the past two years we've started to get a clear picture of how teachers use these various assessment techniques in their classrooms.

In this second part of the study, our job is just as important as the first part. Now, we're trying to get an accurate picture about how much time it all takes, and again we want to get that information directly from administrators and teachers.

Now, I'll get back to this later, but let me mention that just as we are interested in the total range of assessment techniques your teachers use in your classrooms

Anyway, that's the project in a nutshell. Is there anything you'd like me to clarify before we go on?

Let me emphasize that your participation is voluntary and that each person included in the study will remain anonymous.

Any questions about any of that?

Tape Recording

Now, since I don't want to miss any of what you say, or inadvertantly change your words, I'd like your permission to tape record our talk. (No identifying details will appear on the tape label, and only our project staff will be allowed to hear the tape if they need to transcribe it.) If at any point you want to turn the recorder off, you just need to press this button. (DEMONSTRATE)

So is it okay if we tape record?

Let's begin with some background information.

I. Before we start exploring the testing issue, I would like some background information to get an idea of the context in which the testing situations occur.

1. First, I'd like to know how long you've been at this school.

2. Have you held administrative positions elsewhere?
   Probe: Where assigned previously?

3. Are the students in this school grouped in any particular way? (If administrator needs clarification, provide terms such as: low, middle, high, regular, gifted, cross-grade.)

4. How are student grouping decisions made?
   Probe: Based on yearly testing, grades, teacher judgment, parent recommendation.

II. Okay. Thank you. Let me briefly describe how we will proceed. First, I'll ask you about the school-wide testing program. Then, we'll talk about the various costs, monetary and psychological, for you, your staff and students. Then, any other comments would also be helpful.

TEST USE PHASE II

CORE QUESTIONS

1. What kinds of tests are given on a school-wide basis?

2. Could you estimate how much money per child is spent on testing?

3. Does anyone make you give particular tests? Who"

4. Approximately when are these tests given during the year? That is, approximate months or points during the year.

5. How much time and whose time is used in activities before, during and after administration?

Before	During	After
ordering tests	supervision	collecting and preparing, shipping tests
informing parents, teachers/staff	insuring proper test conditions	- having them scored
inservice activities		- drawing up reports
allocation of staff, equipment and facilities		- disseminating results
coordination with district office		- verifying completions

6. How much time and whose time is used in activities before, during and after the administration? (Teacher aide, parent volunteer, clerical)

Before	During	After
test construction	setting-up	scoring
preparing materials	administering	grading
inservice activities	proctoring	reviewing

Probe: We just talked about personnel. Have we covered all categories of personnel that have to adjust their routine schedules to perform test related activities.

OVERALL

7. How much time would you say your teachers spend on testing over the year?

8. Can you list for me for each kind of school-wide test, the materials, facilities, and equipment, used in testing?

9. Do these displace other school related activities (use of spaces, e.g., auditorium, cafeteria, cancelled classes)

10. Is there anyone in your school who could tell us about costs and/or purchases connected with testing? Where could we get comprehensive budget records with regard to testing?

11. Are there particular kinds of tests that cause stress, anxiety or concern to you? To your staff (both teaching and non-teaching personnel), students or parents?

12. OK, you have told us about the different monetary and psychological costs related to testing. Given all of this, is it worth the cost?

13. What tests would you eliminate if it were left up to you?

Interviewer:_____

DATA RECORDING SHEET - ADMINISTRATIVE

I. Background

1. Years at this school_____

2. Administrative positions elsewhere:

_____

_____

_____

3. Grouping:_____

_____

4. How grouping is decided:_____

_____

_____

5. Cost, per child, on testing: _____

II. Go to Test Data sheet.

ADMINISTRATOR

1. Test _____ 2. Who says _____ 4. X per year _____
   Type _____ 3. When _____

	WHAT	WHO (circle as apply)	TIME
BEFORE		AM C A T PV	
		AM C A T PV	
		AM C A T PV	
		AM C A T PV	
		AM C A T PV	
		AM C A T PV	
		AM C A T PV	
		AM C A T PV	
DURING		AM C A T PV	
		AM C A T PV	
		AM C A T PV	
		AM C A T PV	
		AM C A T PV	
		AM C A T PV	
		AM C A T PV	
		AM C A T PV	
AFTER		AM C A T PV	
		AM C A T PV	
		AM C A T PV	
		AM C A T PV	
		AM C A T PV	
		AM C A T PV	
		AM C A T PV	
		AM C A T PV	

AM= Administrator C= Clerical A= Aide T= Teacher S= Student PV= Parent Volunteer

6. Time of Teacher

7 & 8.

Materials	Displace	Facilities	Displace	Equipment	Displace

9. Test costs/purchases

10. Anxiety: Administrative Staff

Anxiety: Teachers

Anxiety: Kids

Anxiety: Parents

11. Is it worth it?

12. Tests to eliminate:

APPENDIX C

PICK UP STICKS

INTERVIEWER _____ DATE _____ SUBJECT # _____

SCHOOL _____ GRADE _____ M F B W H A PI O

PART I: LIKE/DISLIKE SORT

Instructions: I'm going to describe some things that probably happen in school, and, at the same time, I'm going to give you a set of sticks with these activities written on them. I want to get a sense of what you think about these things.

1. I want you to make three piles, sorting the sticks into groups of things that you like to do and things that you don't like to do. In the middle, place things that happen at school that you don't have an opinion about. (Display labels as you speak.)

   Now, within each pile, put them in order of things that you like, with the best or favorite activity on the top. Do the same for the things that you don't like, putting the activity that you hate the most on top.

Things I like		Things in the Middle		Things I don't like	
Activity	Letter	Activity	Letter	Activity	Letter

INSTRUCTIONS WRITTEN ON ATTACHED PAGE

EXAMPLE:	+	7	6	5	4	3	2	1	-
ASSEMBLIES	Fun								Not fun
	Smart								Dumb
	Important								Unimportant
	Calm								Worried
(a) HOMEWORK	Fun								Not fun
	Smart								Dumb
	Important								Unimportant
	Calm								Worried
(b) WRITING A STORY	Fun								Not fun
	Smart								Dumb
	Important								Unimportant
	Calm								Worried
(c) STANDARDIZED TEST	Fun								Not fun
	Smart								Dumb
	Important								Unimportant
	Calm								Worried
(d) P.E. GAMES	Fun								Not fun
	Smart								Dumb
	Important								Unimportant
	Calm								Worried
(e) ANSWERING QUESTIONS	Fun								Not fun
	Smart								Dumb
	Important								Unimportant
	Calm								Worried

(f) RECESS/ NUTRITION	Fun	___	___	___	___	___	___	___	Not fun
	Smart	___	___	___	___	___	___	___	Dumb
	Important	___	___	___	___	___	___	___	Unimportant
	Calm	___	___	___	___	___	___	___	Worried
(g) TAKING A CHAPTER TEST	Fun	___	___	___	___	___	___	___	Not fun
	Smart	___	___	___	___	___	___	___	Dumb
	Important	___	___	___	___	___	___	___	Unimportant
	Calm	___	___	___	___	___	___	___	Worried
(h) TALKING WITH FRIENDS	Fun	___	___	___	___	___	___	___	Not fun
	Smart	___	___	___	___	___	___	___	Dumb
	Important	___	___	___	___	___	___	___	Unimportant
	Calm	___	___	___	___	___	___	___	Worried
(i) TAKING A TEACHER MADE QUIZ	Fun	___	___	___	___	___	___	___	Not fun
	Smart	___	___	___	___	___	___	___	Dumb
	Important	___	___	___	___	___	___	___	Unimportant
	Calm	___	___	___	___	___	___	___	Worried

	HOMEWORK	TEACHERS' QUESTIONS IN CLASS	STANDARD-IZED TESTS	CHAPTER TESTS	TEACHER MADE TESTS
MY TEACHER					
MY FOLKS					
ME					
KIDS IN MY CLASS					